THIN-FILM INTRACORTICAL RECORDING MICROELECTRODES

Quarterly Report #1

(Contract NIH-NINDS-NO1-NS-0-2329)

September – December 2000

Submitted to the

Neural Prosthesis Program
National Institute of Neurological Disorders and Stroke
National Institutes of Health

by the

Center for Wireless Integrated MicroSystems
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan
48109-2122

January 2001

Thin-Film Intracortical Recording Microelectrodes

Summary

The goal of this new contract is to develop a family of active recording probes suitable for fundamental studies in neurophysiology and for use in neural prostheses. The probes have 64 sites, of which eight can be selected for simultaneous use by the external world. On one of the probe designs (PIA-2B/3B), the neural signals are buffered and then passed directly off chip, whereas on the other (PIA-2/-3) the signals are amplified, multiplexed, and then passed off chip to minimize external leads. Both two-dimensional (2D) and three-dimensional (3D) versions of these probes are being developed.

During the past quarter, probes having 16 sites have been chronically implanted in guinea pig auditory cortex in order to study site lifetime as a function of site size and placement. The probes contained a 1000μm^2 site at the tip and a normal 177μm^2 site located in the center of the shank 25μm behind it. After 9 weeks, no significant differences in recording quality have been observed over these sites, although the impedances slowly fluctuate. We have also implanted 32-channel "brain-in-the-box" probes; recorded activity on the first of these implants stopped after two weeks, and we suspect a cable break. Another array has been implanted and is being monitored carefully.

We have also completed a fabrication run containing a large number of active probes. Fifteen different probe designs were fabricated, including 64- and 96-site structures. The front-end-selected Phoenix3 probe has been used to record spontaneous and driven activity from 64 sites on 8 shanks. Some of these active probes contain special test modes that allow a test signal to be generated on-chip and applied to the sites. This feature has been verified for the first time and found to work satisfactorily so that active probes can be tested in-vivo at any time to examine site impedances. The 96-site buffered probes have been used to record from single units in guinea pig auditory cortex and in cerebellum. They are also currently being used for experiments in rat hippocampus at Rutgers University.

The probes fabricated employ a variety of readout circuits including capacitively-coupled probes designed to avoid dc offsets. These probes have measured gains of 36dB, bandwidths from 68Hz to 25kHz, and a power dissipation of 77μW. An eight-channel multiplexed probe has also been fabricated. Optically-induced offsets remain a problem on some of these probes, and the circuitry is being redesigned to eliminate the problem. In an effort to increase the number of sites that can be monitored, an in-vivo spike-detection circuit is being designed. It will be fabricated using the HP 0.5μm process available through MOSIS and mounted on the platform holding the probes. The spike detection threshold an be input by the user so that the location and time of occurrence of

the spikes will be transmitted, eliminating the bandwidth consumed in transmitting baseline noise in full-analog systems. Eventually, the digitized spike amplitude could also be transmitted. The system is being designed to handle 1024 sites in a 3D array with 128 active channels. At a 2.56MHz clock, the time of any spike can be resolved to less than 0.4µsec.

Work has also gone forward with the design of a wireless interface for the probes. A 10b sigma-delta analog-to-digital converter (ADC) has been designed to achieve a Nyquist sampling rate of 20kHz. The dynamic range is 62dB. Simulations of the operational transconductance amplifier show a power dissipation of 1.7mW, a dc gain of 120dB, a unity-gain frequency of 28MHz and a dynamic range of 69dB. A two-phase non-overlapping clock has also been designed for this system with a simulated power dissipation of 226µW. A decimator has also been designed. The ADC has been laid out and signal and data transmission protocols for the wireless system have been defined. The system will use an amplitude-shift keyed format. The goal is a fully wireless recording system by the end of summer.

Thin-Film Intracortical Recording Microelectrodes

1. Introduction

The goal of this program is the realization of batch-fabricated recording electrode arrays capable of accurately sampling single-unit neural activity throughout of volume of cortical tissue on a chronic basis. Such arrays will constitute an important advance in instrumentation for the study of information processing in neural structures and should be valuable for a number of next-generation closed-loop neural prostheses, where stimuli must be conditioned on the response of the physiological system.

The approach taken in this research involves the use of solid-state process technology to realize probes in which a precisely-etched silicon substrate supports an array of thin-film conductors insulated above and below by deposited dielectrics. Openings in the dielectrics, produced using photolithography, form recording sites which permit recording from single neurons on a highly-selective basis. The fabrication processes for both passive and active (containing signal-processing circuitry) probe structures have been reported in the past along with scaling limits and the results of numerous acute experiments using passive probes in animals. In moving to chronic implant applications, the major problems are associated with the preserving the viability of the sites in-vivo (preventing tissue encapsulation of the sites) and with the probe output leads, both in terms of their number and their insulation. The probe must float in the tissue with minimal tethering forces, limiting the number of leads to a few at most. The encapsulation of these leads must offer adequate protection for the megohm impedance levels of the sites while maintaining mechanical lead flexibility.

Our solution to the lead problem has involved two steps. The first has been to embed circuitry in the probe substrate to amplify and buffer the signals and to multiplex them onto a common output line. Using this approach, signal levels are increased by factors of over 100, impedance levels are reduced by three to four orders of magnitude, and the probe requires only a few leads for operation, independent of the number of recording sites. A high-yield merged process permitting the integration of CMOS circuitry on the probe has been developed, and this circuitry has been designed and characterized. The second step has involved the development of silicon-based ribbon cables, realized using the same probe technology, to conduct the neural signals to the outside world. These cables have shown significant advantages over discrete leads, both in terms of the ease with which chronic implants can be assembled and in terms of the ability of the cables to survive long-term biased soaks in saline. The cables can be built directly into the probes so that they come off of the wafer as a single unit, requiring no joining or bonding operations between them. The cables are also significantly more flexible than previously-used discrete wire interconnects.

This new contract calls for the development of active probes for neural recording. A 64-site 8-channel probe with site selection and signal buffering but no multiplexing has

been developed (PIA-2B) along with a high-end multiplexed probe that includes gain (PIA-2). These probes are now being refined and applied to in-vivo applications. Investigations are on-going to better understand site encapsulation, which limits the lifetime of chronic recording structures, and telemetry is being developed to allow the probes to be operated over a wireless link, eliminating the percutaneous connector.

During the past quarter, we have continued to study recording lifetime as a function of site size and placement. A fabrication run of active probes has been completed and characterized, including probes that are ac-coupled using capacitors to set the gain. Work has also gone forward on a completely wireless probe system with the design of a sigma-delta-based analog-to-digital converter. Work in these areas is discussed in the sections below.

2. Passive Probe Development

Understanding and improving chronic electrode lifetime remains a major focus for the new contract. While chronically implanted probes sometimes maintain activity for periods of months, more typically we see a degradation in signal characteristics over the first few weeks. This is exhibited as an increase in impedance in parallel with a decline in the signal-to-noise ratio of neural activity. It has been hypothesized that the degradation in recording characteristics is due to protein fouling of the surface and/or glial encapsulation. Other electrode types, including microwires and the Utah array, have demonstrated somewhat longer recording lifetimes, although they are not immune from these effects. It has been speculated that these electrode types record over longer periods because their recording sites are larger and are at the tips of the shanks, subjecting them to more relative movement due to natural brain pulsations. This movement may aid in the continual break down of the layer of encapsulation that obstructs the recording sites. Several probes were recently designed with sites at the shank tips and sites that extend beyond the side of the substrate in an effort to mimic this type of electrode structure (Fig. 1).

A probe with tip sites (Fig. 1) was implanted in guinea pig auditory cortex. Each of the 8 shanks on this design has a large recording site ($1000\mu m^2$) at the tip and a standard recording site ($177\mu m^2$) 25µm back in the center of the shank. Driven unit activity has been observed on at least one of the electrode sites throughout the entire 9-week implant period. The active channels have changed from week to week, however, and definitive statements cannot yet be made about the recording quality of tip versus middle sites. In fact, both site types have shown comparable signal quality at some times during the implant. Figure 2 shows an example of recordings obtained from two recording sessions, 8 and 9 weeks post implantation. Activity can be seen on both tip sites (odd numbered sites) and center sites (even numbered sites). Plots of 1kHz impedance over time are shown in Fig. 3. Impedances have fluctuated for both site types but for the most part have stayed within the range of impedances that are capable of recording activity. We will continue to monitor this implant as long as it remains viable.

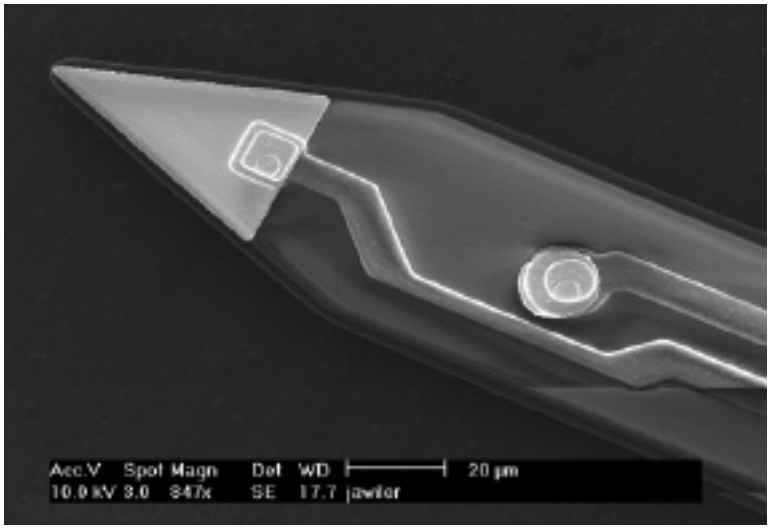


Fig. 1: SEM of one shank of a new 8-shank probe designed to investigate effects of site placement on chronic recording lifetime. In addition to a typical recording site in the middle of the shank, the probe has a larger site at the tip of the shank.

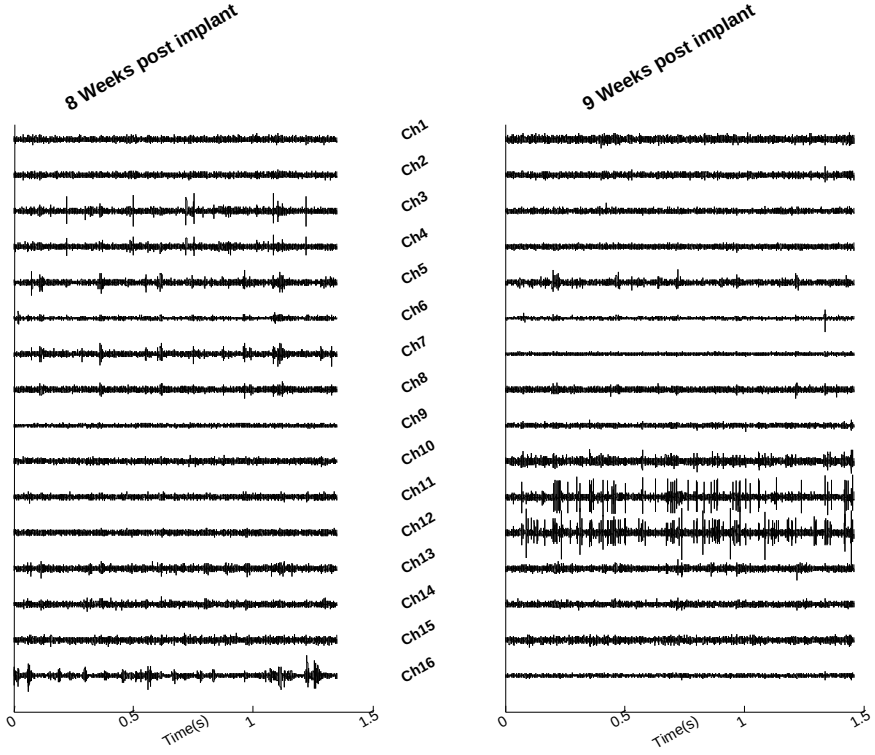


Fig. 2: Recordings from guinea pig auditory cortex 8 and 9 weeks post implantation. The odd numbered channels are the tip sites.

The first fully-wired 32 channel "Brain in the Box" (BIB) probes were implanted this past quarter. The BIB is a 3-D structure constructed from a single cable that forks into two separate probes that are unfolded and inserted into slots in a platform. This results in a structure that has the two probes having sites that face one another. A photograph of the device and its packaging is shown in Fig. 4. Two 18-pin Omnetics NANO connectors are used to provide percutaneous access to the signals. The first of the two BIB implants had discriminable cells on 15 of the 32 channels for two weeks, after which time the impedances abruptly increased. A cable break was suspected and was verified upon sacrificing the animal. Cable breakage has been a problem in past BIB implants. These probes are designed to maintain a low cable profile as they exit the skull defect (i.e., the cable should lie almost flat to the skull). Breakage has probably been partially due to the cable being too long, forcing precarious maneuvering and placement during closure. In addition, the cable comes straight off the back of the platform. This is non-optimal since the thickness of the skull must be cleared before the cable enters the connector. Future designs may incorporate an angled cable to permit the cable to clear the skull edge without causing torque on the implanted probe. We are also looking at coating the entire cable with a thin layer of silicone rubber to protect the cable as it leaves the cranial defect. A second BIB was very recently implanted and will be monitored in the coming months.

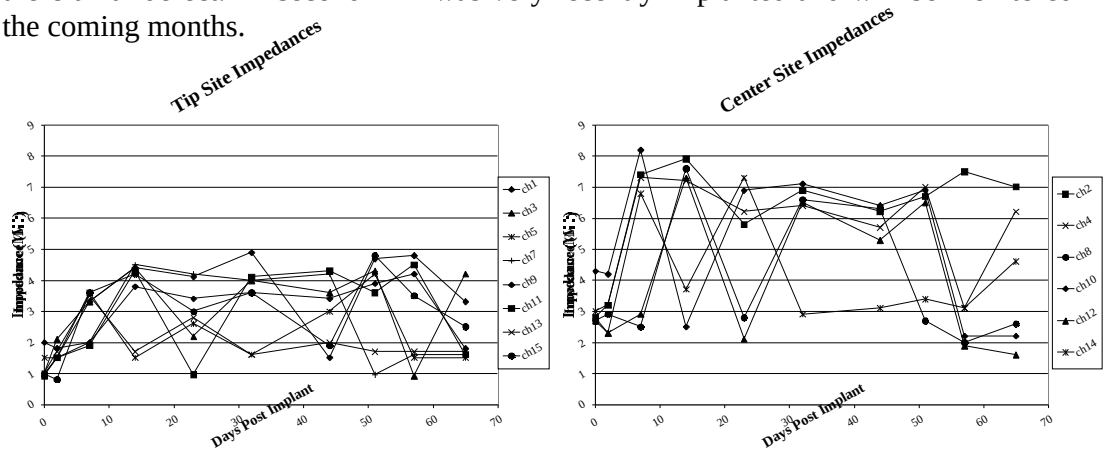


Fig. 3: Plots of 1kHz impedance over time for a chronically implanted probe (Fig. 1). While fluctuations are evident, many sites remain within the range of impedances that are capable of recording unit activity. The tip sites impedances are lower due to the larger site area.

Immunohistochemistry:

During this past quarter, unconnected four-shank electrodes were implanted into the occipital cortex of three guinea pigs. After three weeks, the guinea pigs were sacrificed and the tissue harvested for immunohistochemical evaluation. The tissue is currently in process. Our goal is to identify the tissue/cellular components immediately surrounding the implanted electrode to better understand why chronic electrodes fail over time.

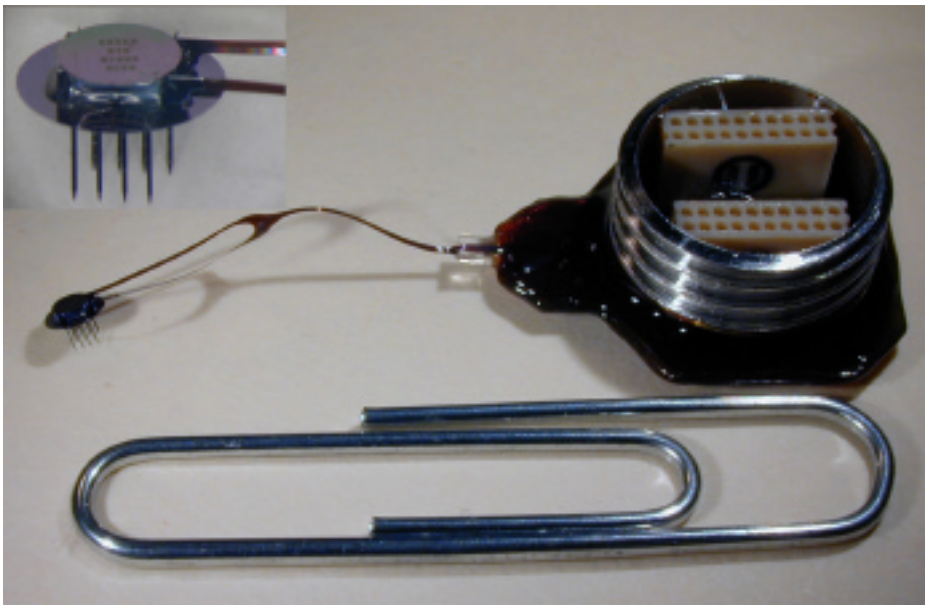


Fig. 4: A fully-connected 32-channel BIB probe. Each of the 8 shanks has 4 sites.

3. Active Recording Probe Development

During the past quarter, an active probe fabrication run was completed. Fifteen different probes were fabricated, as well as test chips for verification of new circuit designs. Several of the probe designs have been described in detail in previous quarterly reports; a summary is given in Table 1.

In order to increase yield, the probes were separated into two mask sets and processed in parallel. One mask set contained most of the larger probes along with the probes that required gold electroplating. This allowed us to process more wafers for the larger probes, which typically experience lower yield. In addition, we were able to release most of the probes for testing before electroplating of the chronic probes was completed. The fabrication run was completed in ten weeks with excellent probe yield and close to expected process parameters.

(a) Front-end-selected Probes

In-vivo Recordings

Both PIA-2B and the Phoenix probes are front-end-selected probes that connect 8 out of 64 sites onto buffered output lines at any given time. The time required for reconfiguring the site selection is relatively small (less than 1msec), so that a large area of tissue can be quickly explored for unit activity.

Probe Name	***Purpose***	***Testing Status/Results***
PIA-2B	64 site front-end-selected, buffered probe	Benchtop and in-vitro testing begun. Functioning as designed.
Phoenix 1-3	64 site front-end-selected probes with improved buffers/amplifiers and DC stabilization	Functioning well on benchtop and in-vitro. In-vivo testing begun with good results.
DC1-4	small probes for comparing DC stabilization techniques	No testing begun.
Probe_2_2_1 Probe_4_4_1 Probe_4_4_2	four shank, four site probes with newly designed closed-loop amplifier in different feedback configurations	Tested on benchtop and in-vitro. High light sensitivity which should be improved with addition of gold shield.
Buzsaki-96	96 site buffered probe	Functioning as designed on benchtop, in-vitro and in-vivo. Being used in animal experiments at Rutgers.
Buzsaki-64	64 site buffered probe	Not functioning well- more benchtop testing necessary to isolate problem.
Phoenix_chronic	front-end selected probe for chronic implantation	No testing begun.
Passage	front-end selected probe with closely spaced sites for demonstration of "electronic passage."	No testing begun.

Table 1: Recently-fabricated active probe designs.

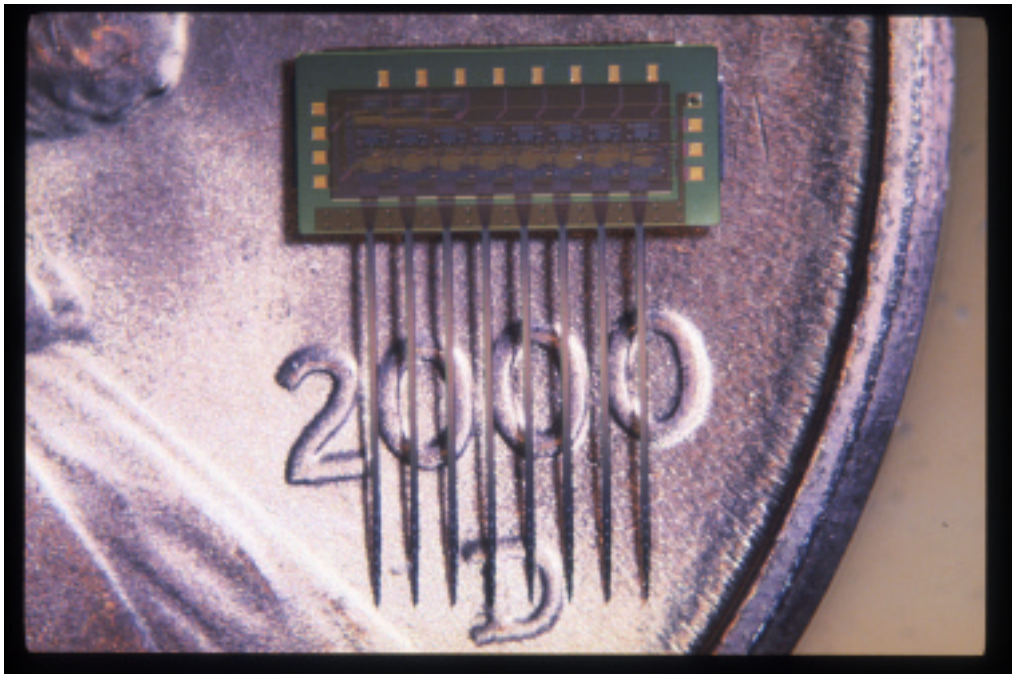


Fig. 5: The 64-site eight-channel probe Phoenix3.

In the past quarter, Phoenix3 has been used to record spontaneous and driven activity from 64 sites on 8 shanks (Fig. 6). This probe is designed to perform a depth scan, meaning that a digital pulse is sent to the probe to select the next group of eight sites drawn one from each shank at a given depth. Roughly 400msec of data recorded in-vivo from guinea pig auditory cortex are aligned with the corresponding site from which the data were obtained. 100msec white noise bursts were presented to the contralateral ear during the recording interval. Driven activity is clearly evident on several sites on shanks five and eight (counting from the left), which can be seen in the periodic nature of the large spikes present.

It is important to note that the data in this figure were not collected simultaneously; that is, traces at a given depth correspond to simultaneous data while those at different depths correspond to successive trials. The underlying assumption is that an accurate picture of neural activity can be assembled by performing a sequential scan.

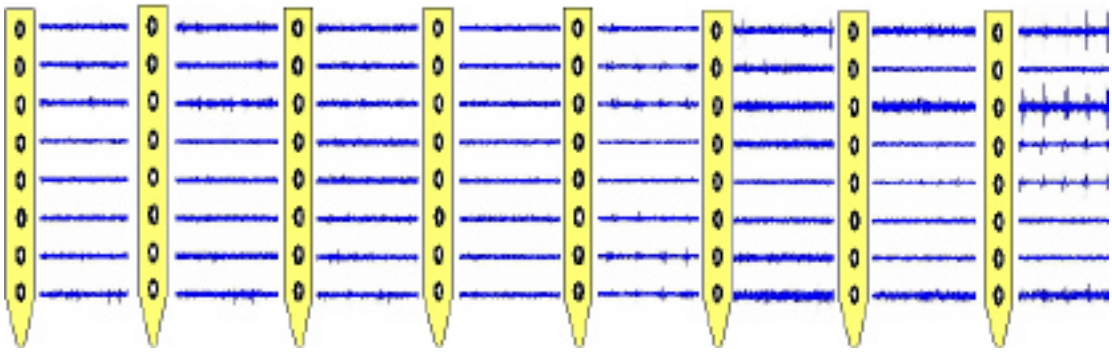


Fig. 6: In-vivo recording from front-end-selected probe Phoenix3.

The recordings from shank eight are given in greater detail and a larger time extent in Fig. 7. Driven activity is clearly visible on sites one, three, four, five, and eight. Detail of a single 100msec stimulus presentation (including vertical scale) recorded from site eight of shank eight is given in Fig. 8. A peri-stimulus time histogram in Fig. 9 clearly shows that the arrival time of spikes is highly correlated to the onset of sound stimulus.

The data from these experiments verify that high-quality recordings can be obtained from a large number of sites with few output leads. The data presented, however, raise some interesting questions which need to be addressed in the coming months. For example, how many cells are involved in producing the spikes evident on sites 3, 4 and 5 of shank eight? (see Fig. 7) To answer this question, we would like to have simultaneous recordings from these sites. Using our knowledge of cell latencies, we could then take a detailed look at spike arrival times to hypothesize how many independent cells are firing. Phoenix3 was used in this experiment because of the simplicity of its digital interface; however, its design does not allow simultaneous sampling from all sites on a single shank. PIA-2b allows us this flexibility, in addition to the ability to select virtually any combination of adjacent sites. Work on testing and

verification of PIA-2b is ongoing, and an external user interface for probe use and configuration is being developed. Experiments using PIA-2b are planned in the coming quarter.

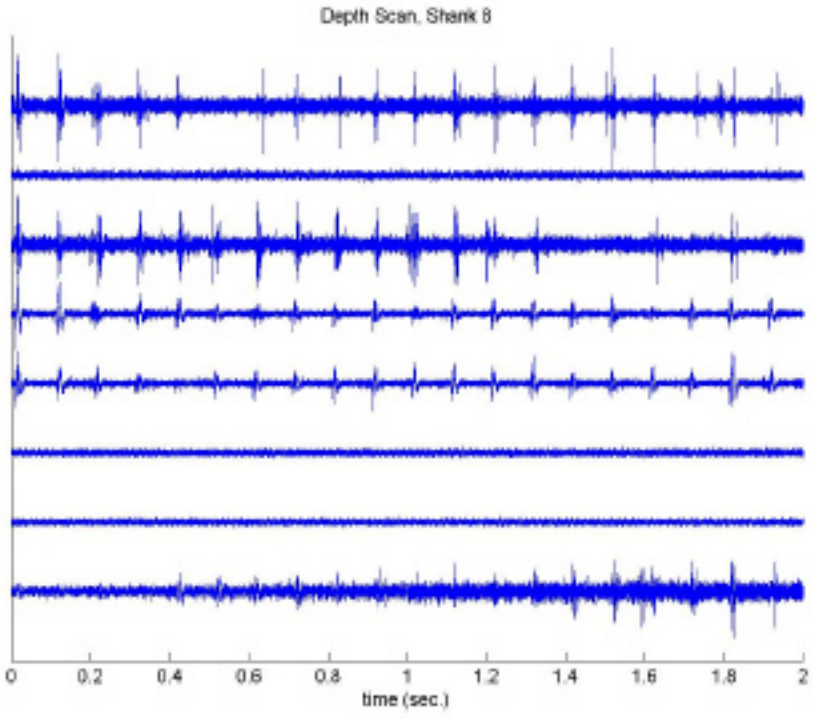


Fig. 7: Driven activity recorded on a single shank of Phoenix3.

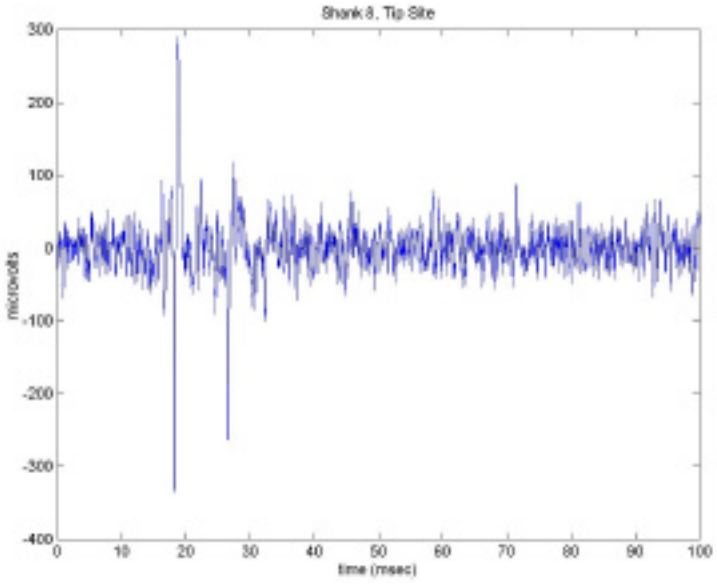


Fig. 8: Two spikes recorded by Phoenix3 during a single 100msec white noise burst presentation.

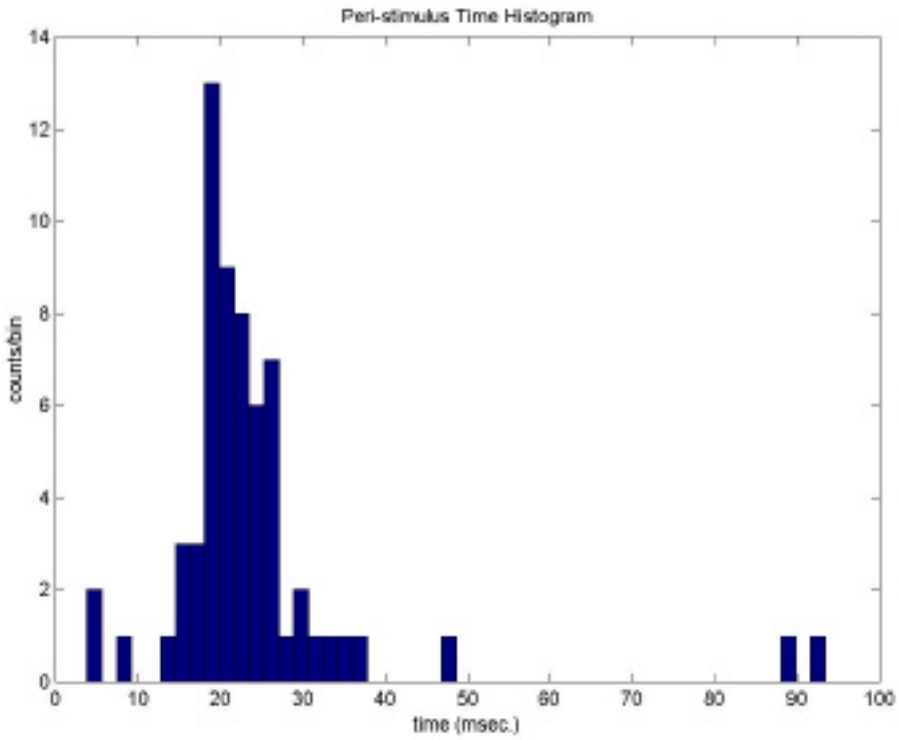


Fig. 9: Peri-stimulus time histogram for shank 8, site 8 clearly showing driven activity.

Also interesting is the gradual appearance of driven activity on site eight. It is likely that this was caused by movement of the probe during the recording interval. Such movement should correspond to a similar change in spike magnitude on other sites. However, since there was little driven activity on the other shanks at this depth during this time, it is difficult to check this hypothesis. As described above, a front-end-selected probe has been designed with closely spaced sites. An experiment is planned with this probe during which the probe will be advanced in small steps between successive recordings from all sites. In this way, we hope to illustrate that electronic site selection can be used to re-center a cell for maximum spike amplitude after a probe movement.

Probe Test Modes

Several test modes have been built into the front-end-selected probes in order to allow for self-test and characterization features. Testing to date has focused primarily on the site impedance measurement capabilities of the probes. In the site impedance mode, a test signal is coupled into the site through an on-chip capacitor and transistor. The output of the buffered channel is measured, and from the magnitude and phase of the output signal relative to the input, we can deduce the magnitude and phase of the site impedance. It is important to note that this measurement would not be possible without

the on-chip buffers, which reduce the crosstalk between the signal input channel and the output channel to acceptable levels.

Figure 10 illustrates the results of a benchtop verification of the site impedance test mode. In order to conduct this test, a probe was bonded to a 24-pin dip socket. In addition to all of the power and I/O lines, several of the sites were bonded directly out to pins. We were then able to connect capacitors of known value between the site and ground. By putting the probe in site impedance mode, we were able to compare the measured impedances with the known values (at 1kHz). As can be seen from the graph, there is fairly good agreement between the measured and expected values of impedance. Similar results were obtained using purely resistive test impedances, and combinations of capacitive and resistive impedances.

We were able to measure site impedances in saline using the site impedance test mode. The values of 1.5 to 2MΩ correspond well to the values typically measured for unactivated irridium sites of this type. The presence of the grounded substrate electrode in the saline makes it difficult to get accurate four-point measurements of the type usually made for site impedances on passive probes. While this makes direct validation of the measurement difficult, it is also part of the reason why this functionality was included into the probe in the first place.

Fig. 10: Measured vs. expected impedance using discrete capacitors for calibration.

3.2 96- and 64-Site Buffered Probes

Both in-vitro and in-vivo testing of the high-channel-count buffered probes have been carried out. These probes are designed for use in experiments where simultaneous recordings from many channels are desired. The on-chip buffering eliminates the need for a discrete component unity-gain headstage amplifier, which can be extremely cumbersome with a large number of channels.

The 96-site probe (Fig. 11) has been used successfully to record single units in both guinea pig auditory cortex and cerebellum (Fig. 12). The fully-fabricated probes typically have between 90 and 94 functional recording channels, with the most common failure mode being an open circuit in the relatively long polysilicon leads.

This probe is currently being tested and used in experiments in rat hippocampus in Gyuri Buzsaki's laboratory at Rutgers. A passive version of the probe has also been fabricated for use in chronic behavioral experiments, and chronic data has been obtained from four animals to date. These experiments stress the importance of developing a better understanding of the chronic recording capabilities of the active probes, so that these probes can be utilized in chronic preparations in the future. A series of chronic in-vitro tests of active probes are planned for the coming quarter.

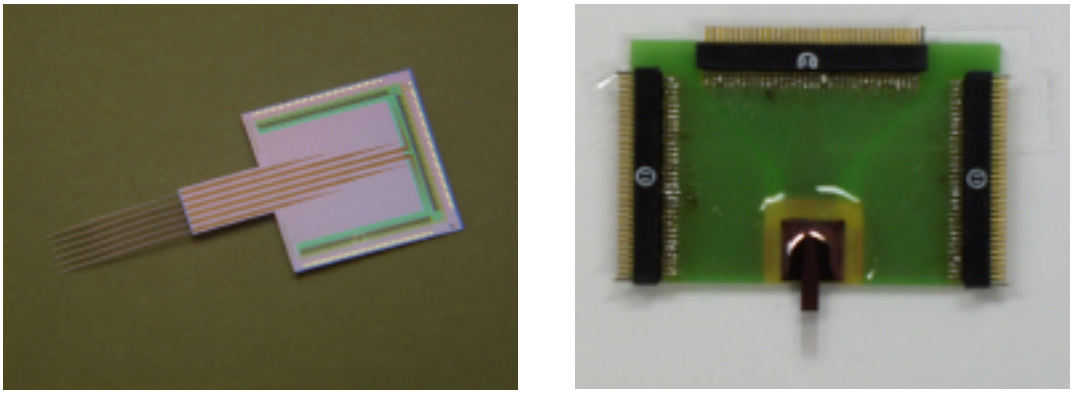


Fig. 11: A 96-site buffered probe, fully bonded at right.

4. Readout Circuitry for Active Recording Probes

Recently many new active probes and electronics have been fabricated and tested. Figure 13 shows a four-channel amplified probe. A schematic of the amplifier feedback configuration on this probe is shown in Fig. 14. This amplifier has been tested on the die level, and the ac response of the amp is shown in Fig. 15, with its performance characteristics summarized in Table 2.

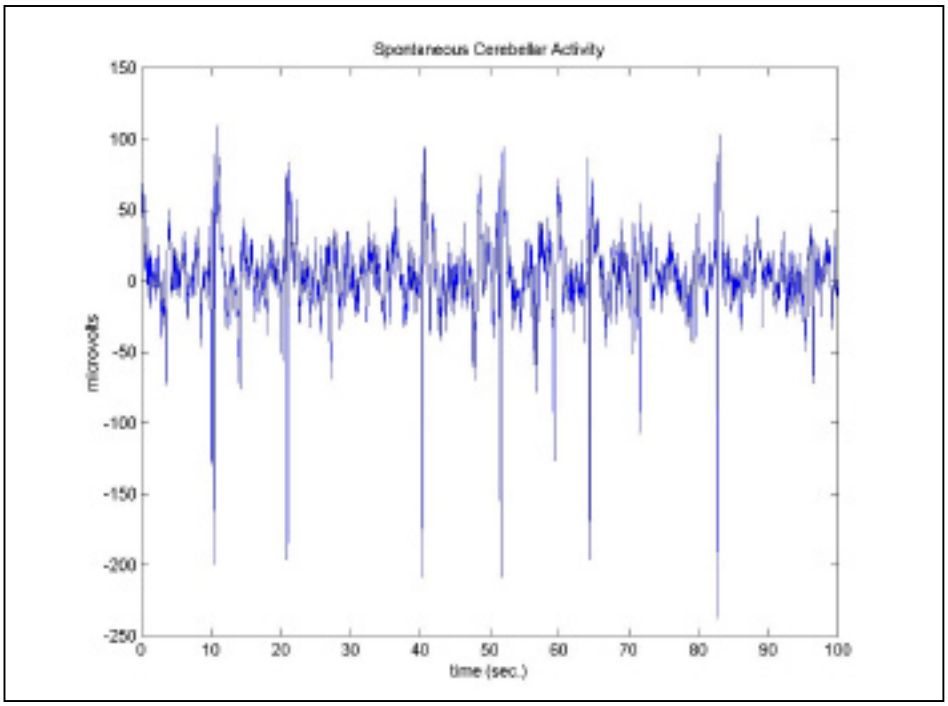


Fig. 12: Spontaneous activity in guinea pig cerebellum recorded with a 96-site buffered probe.

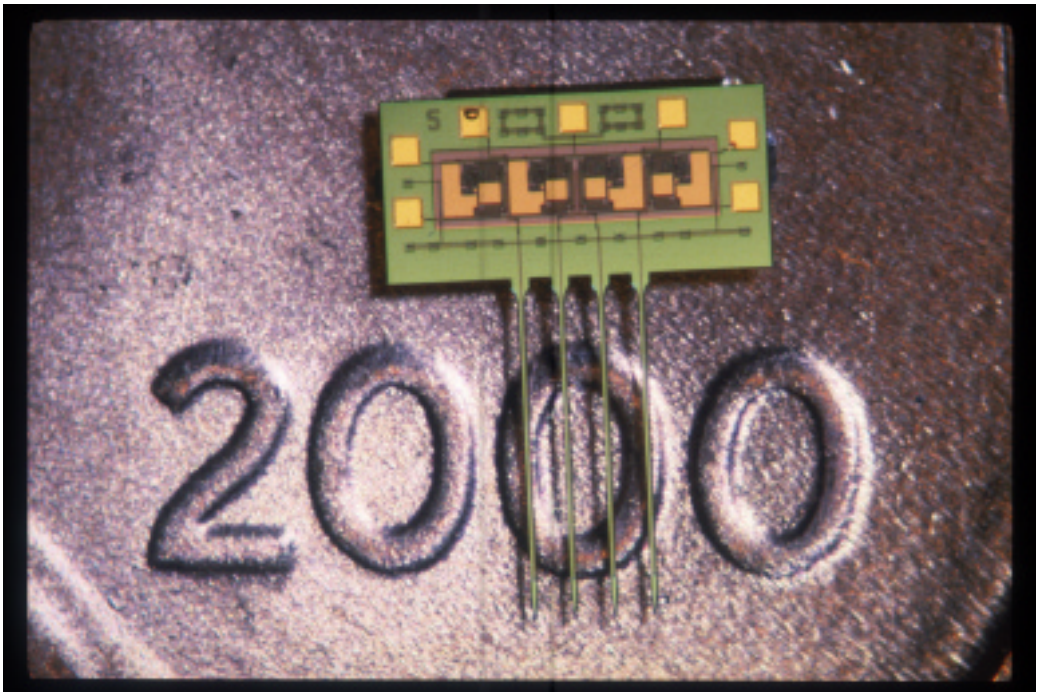


Fig. 13: Four -channel amplified probe

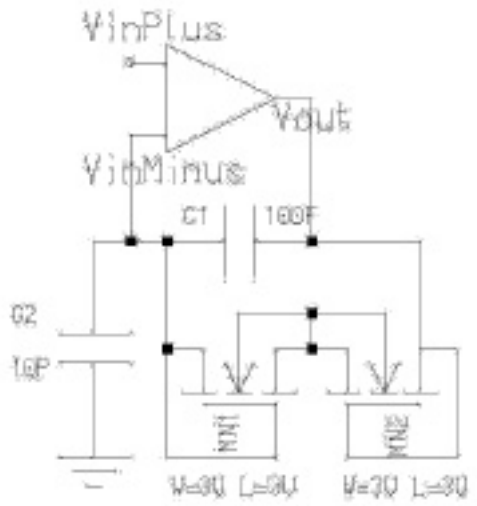


Fig. 14: Amp1 schematic

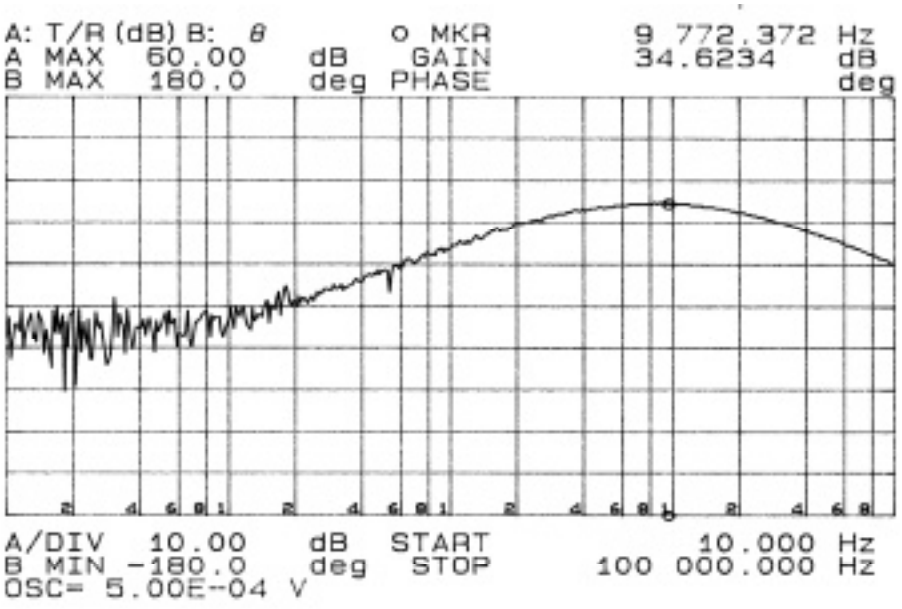


Fig. 15: Gain vs. frequency for the amplifier in Fig. 14

	Simulated	*Measured*
Gain	38.6db	34.6db
Bandwidth	1.3Hz to 15KHz	3Hz to 20KHz
Power Consumption	95.5µW	76.8µW
Input Referred Noise	5.4µVrms/√Hz	--------
Layout Area	0.086mm^2	0.086mm^2

Table 2: Performance specifications of amplifier in Fig. 2.

Another four-channel amplified probe has been fabricated with the amplifiers connected in the feedback configuration shown in Fig. 16. This amplifier has also been tested at the die level and the ac response of the amp is shown in Fig. 18. The response of the amplifier to a 7mV sine wave at 1kHz is shown in Fig. 17. Table 5 summarizes the performance characteristics of the amplifier.

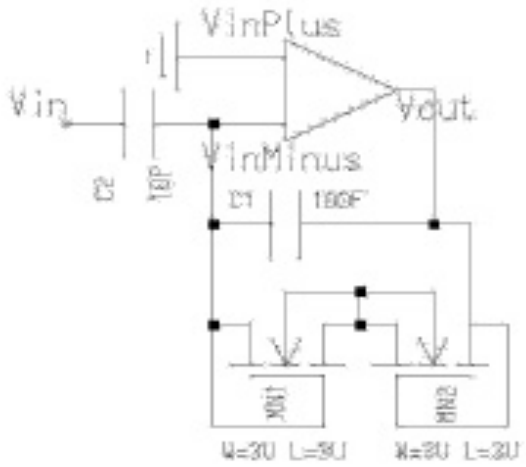


Fig. 16: Amp2 schematic

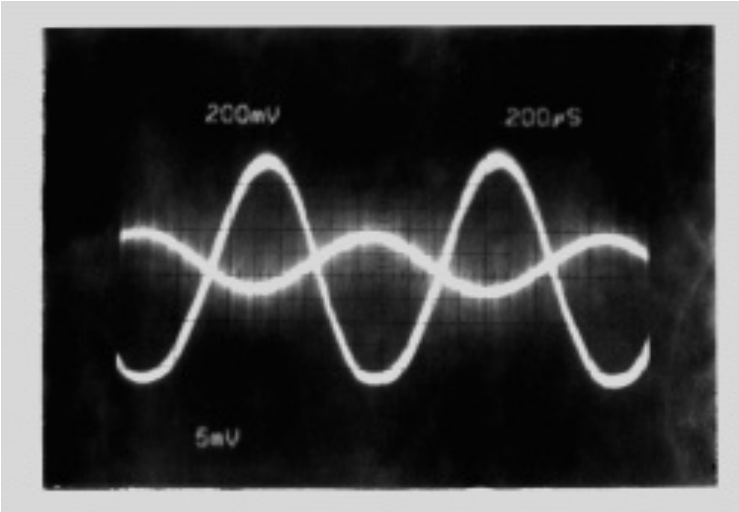


Fig. 17: Measured transient response of Amp2

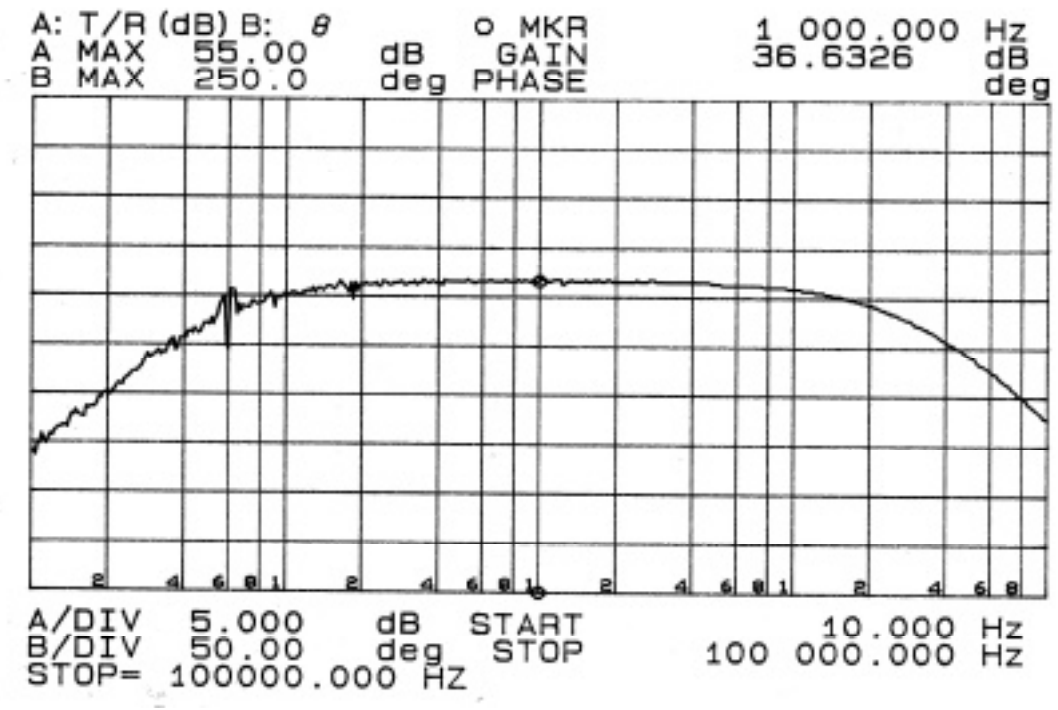


Fig. 18: Gain vs. frequency for the amplifier in Fig. 16.

	Simulated	*Measured*
Gain	36.3db	36.6db
Bandwidth	1.7Hz to 16KHz	68Hz to 25KHz
Power Consumption	95.5µW	76.8µW
Input Referred Noise	6.4µVrms/√Hz	--------
Layout Area	0.078mm^2	0.078mm^2

Table 3: Performance specifications for the amplifier in Fig. 16.

An eight channel, amplified, time-division multiplexed probe has been fabricated and is shown in Fig. 19, with a system level block diagram of the electronics shown in Fig. 20. This probe has been tested on a probe station and the time-division multiplexor is fully functional. The eight pre-amplifiers, connected as shown in Fig. 21, also perform as designed. The output amplifier, however, is not functional. This will be discussed below.

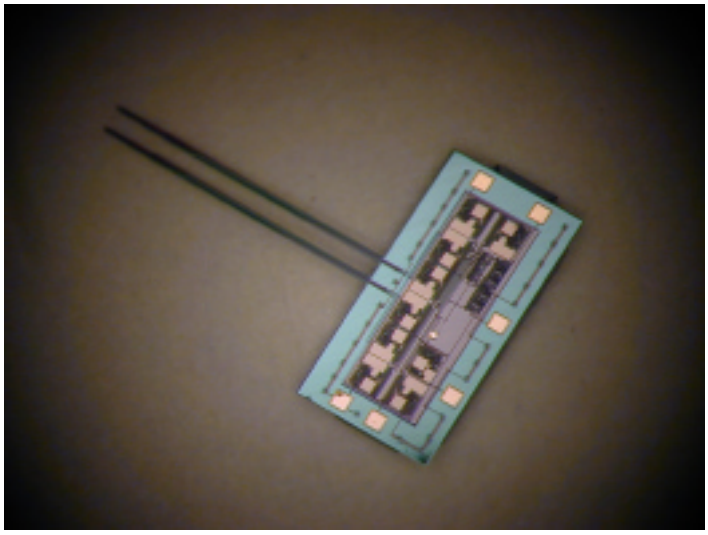


Fig. 19: An amplified multiplexed probe

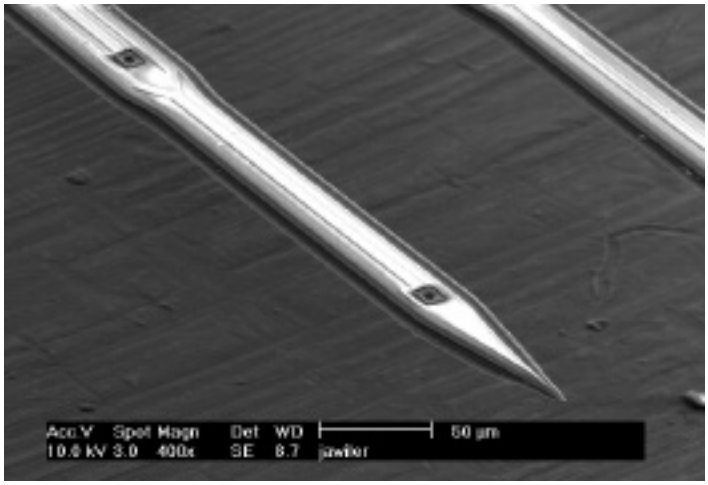


Fig. 20: An SEM of the shank tips

Optical interference is a major problem with all of the amplifier designs. This is especially true after the EDP release, where the circuit areas are thinned to between 40 and 85µm . Before the EDP release, very high intensity light will move the bias points of the amplifiers, but moderate light levels do not significantly effect the amplifier performance. After EDP release, however, the bias points of the amplifier become very optically sensitive even at low light levels. Optical noise also couples into the output signal of the amplifier, drowning out low level neural pulses. A metal shield over the circuit areas will be incorporated on the next processing run to help alleviate this problem. In addition, a design change to the amplifier feedback configuration will be made. Figure 22 shows the current implementation of the diode-connected transistors in the amplifier feedback loop. There is a large parasitic diode between the n-epi and the

floating p-well. This diode has a large depletion area due to the relatively low doping densities in the two regions. This makes the n-epi:p-well diode a very good photo detector. When light hits the diode it creates electron hole pairs that get swept across the depletion layer. This current then flows into the floating p-well and the diode connected transistors. The majority of this optically induced current flows into the output of the amplifier, driving it to the positive supply rail. The current also lowers the resistance of the back biased diode connected transistors, changing the frequency response of the amplifier. Future designs will feature a grounded rather than floating p-well. This will cause any optically induced currents to flow to ground instead of into the diode-connected transistors.

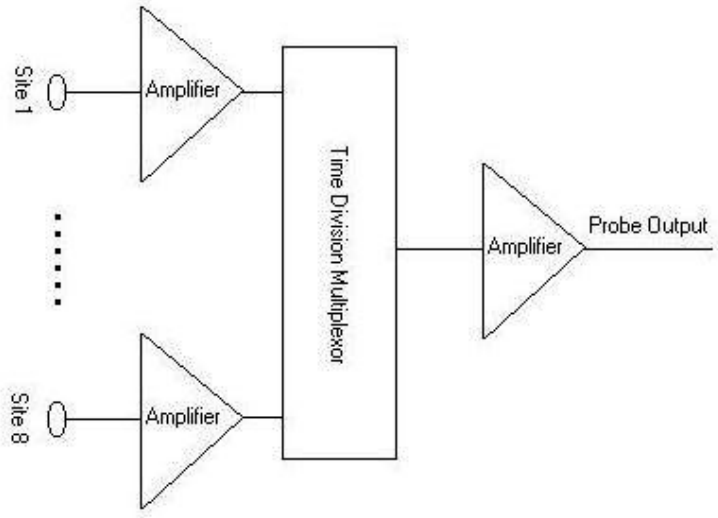


Fig. 21: Block diagram of an amplified time-division-multiplexed probe

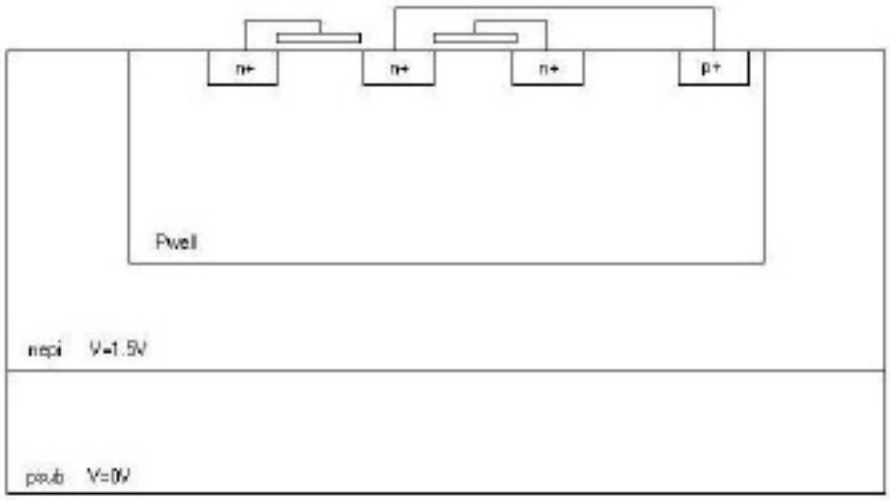


Fig. 22: Implementation of diode-connected transistors

In an effort to increase the number of sites monitored simultaneously without significantly increasing the lead count, a spike-detection ASIC is being designed. The ASIC will be fabricated in the Hewlett-Packard 0.5μm process offered through MOSIS. A block diagram of the spike-detection unit is shown in Fig. 23. The proposed ASIC would be designed to sit on the platform of a 1024-site 3D neural recording array. A front-end selector would condense the number of sites being actively monitored to 128, with 16 active sites on each 2D probe. These sites would be amplified and time division multiplexed on the 2D probes to reduce the number of beam leads required to transfer signals from the 2D probes to the platform. The clocks required to perform the time division multiplexing would be provided by the ASIC. The ASIC itself would have as its inputs the time-division-multiplexed neural signals from the 2D probes. Voltage comparators will be used to determine if there is a neural spike at a site. The thresholds for these comparators will be set by the user. Once a spike has been detected, the address of the site where the spike occurred is put in the output register for serial transmission off of the platform. If the output register is occupied by another site address, the current spike address will be put into one of the eight hold registers and transmitted off chip when the output register becomes free. If all of the hold registers are full, the spike is thrown out and the threshold level of the comparators should most likely be increased. At the proposed clock frequency of 2.56MHz, the response at any site can be resolved to within 390ns. One can clearly imagine that future versions of signal processors might quantize the amplitude of the spike once it has been detected. In such a system, the user would have the amplitude and frequency response data from any spike that exceeded the threshold voltage. The only information lost in this exchange is the noise information at the sites. In many applications and sorting algorithms the background noise is not used and in these circumstances this ASIC could be a powerful tool because it greatly increases the number of sites that can be monitored simultaneously by eliminating the neural noise before transmitting data off of the platform.

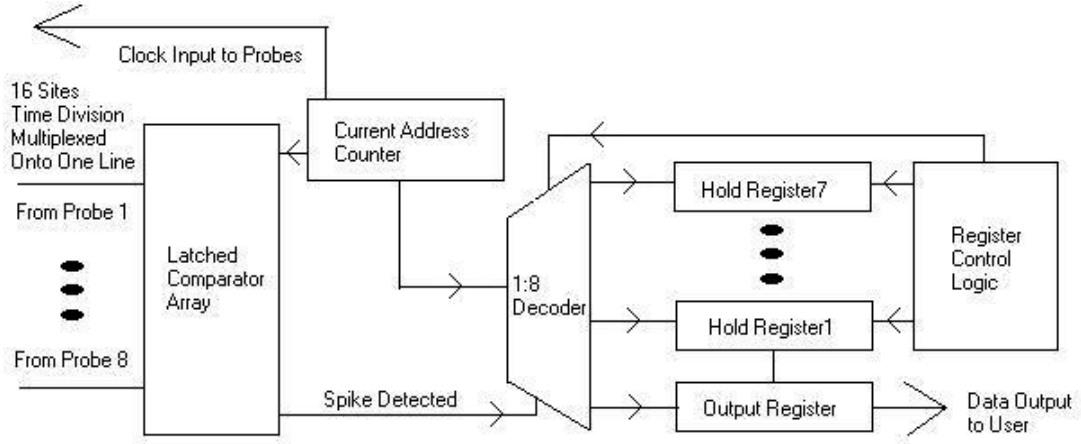


Fig. 23: Spike-detection unit block diagram

5. Design of a Wireless Telemetry Platform for Multichannel Microprobes

As reported in the last progress report, we have been designing a circuit chip to operate multi-channel recording microprobes using a wireless link. This wireless link receives a modulated RF signal from an outside transmitter, decodes data modulated onto this RF carrier, generates low-noise power supplies, and transmits recorded neural data to the outside world. We have reported the design of the front-end circuits needed for power regulation, data recovery in previous reports, and have discussed the design of the analog-digital converter needed to convert analog neural data into a digital format. This section discusses progress made during the past six months in a number of areas, including redesign of the front-end circuits, design and simulation of the sigma-delta analog-digital converter, and layout of a modified front-end circuit.

5.1 Design of the Second-Order Sigma-Delta Modulator

Our goal is to design a sigma-delta A/D converter which achieves 10-bit resolution at a Nyquist rate of 20kHz with minimum power consumption.

Signal Noise Ratio

If we assume the ADC is ideal, the width of each code step is identical and has an equal probability. For an n-bit ADC, the signal to quantization noise ratio, or dynamic range, can be shown as

$$SNR(dB) = 6.02n + 1.76dB$$

Our design requirement of 10 bits corresponds to a dynamic range of 62dB.

An approximate expression for the quantization noise when the quantizer is modeled by an additive white-noise source is

$$SNR(dB) = \frac{3}{2}\frac{2L+1}{\pi^{2L}}M^{2L+1}, (M >> 1)$$

where L is the order of the sigma delta modulator, and M is the oversampling ratio.

Note that if L increases, SNR can be improved for the same oversampling ratio, therefore, higher order sigma delta modulators are preferred. However, modulators higher than second order require careful design to ensure stability. Since our requirement for word length of 10 bits is not very demanding, we choose the order to be two and we checked its validity from its noise, required oversampling frequency and necessary

settling time of the integrator . The block diagram of a second order sigma delta modulator is illustrated in Fig. 24.

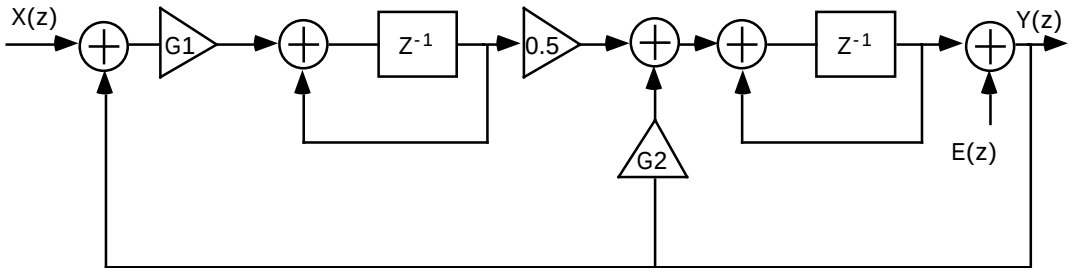


Fig. 24: The structure of a scaled second order modulator

In the z domain, the output of this second order modulator can be written as

$$Y(z) = \frac{0.5G_1 \cdot X(z)\, z^{-2} + E(z)\cdot (1-z^{-1})^2}{(1-z^{-1})^2 + G_2 \cdot z^{-1}(1-z^{-1}) + 0.5G_1 \cdot z^{-2}}$$

For the second order modulator, L=2. If the performance of the modulator is only limited by quantization noise, dynamic range of 62dB can be achieved with an oversampling ratio of M=29. After taking circuit non-ideality into account, a sampling ratio M=64 was finally chosen, which makes the oversampling ratio to be a power of 2 to simplify the subsequent decimation.

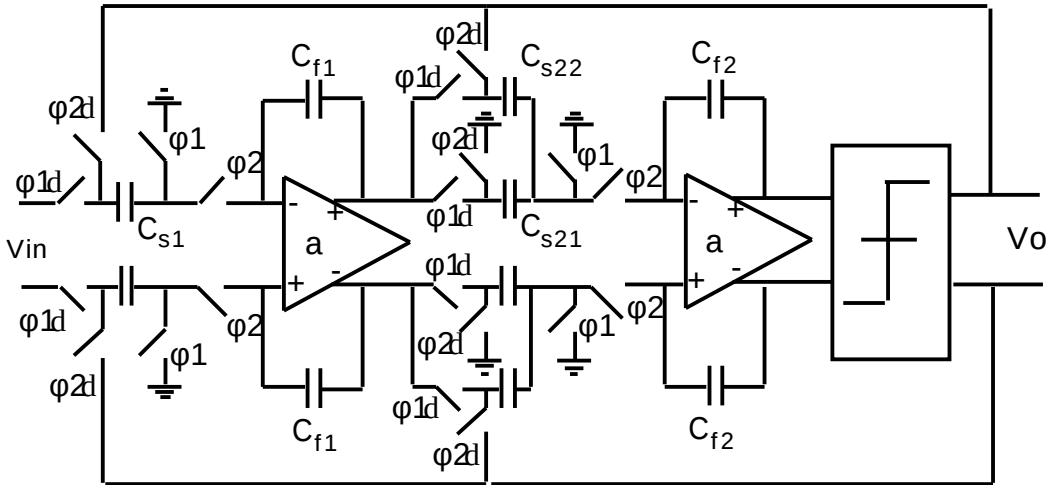
If the Nyquist rate of the neural signal is 20kHz, then the oversampling frequency should be larger than 1.28MHz. Considering the on-chip clock is recovered from an RF carrier of 4MHz, for convenience we choose the oversampling frequency f_s=2.0MHz. In the previous analysis, we ignored the non-ideality of the circuits. In the following, we will account for the thermal noise of the sampling capacitance Cs in the first integrator; then, the SNR can be written as:

$$SNR_{(dB)} \approx \frac{\frac{\Delta^2}{8}}{\frac{\Delta^2}{8}\frac{2\pi^4}{35M^5} + \frac{4kT}{MC_s}}$$

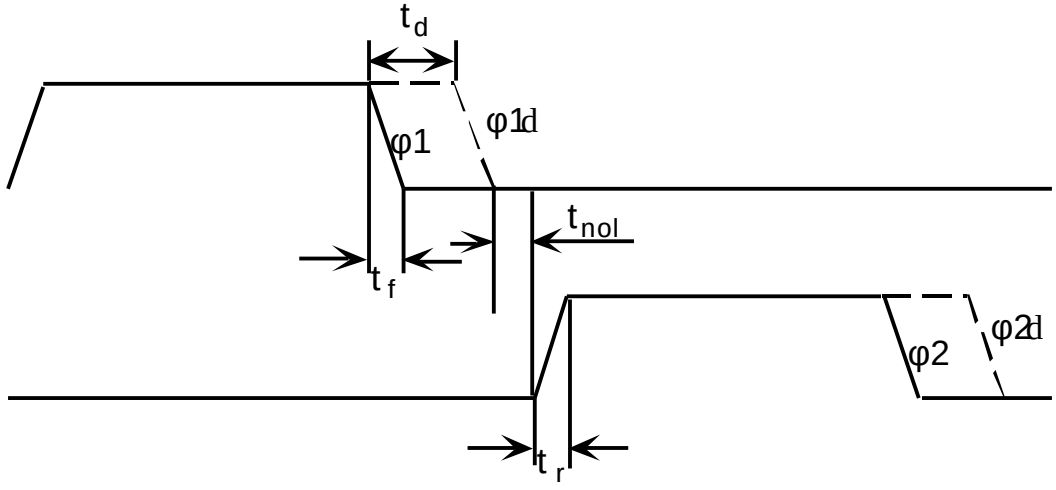
where Δ is 10 volts, because we have the supply voltage of 5V and we will use a fully-differential Operational Transconductance Amplifier (OTA). Therefore, C_s>39fF. MIDAS3, contributed by Stanford University, is used to simulate the switched-capacitor circuits. The settling time of the OTA used in the integrator is then determined mainly by oversampling frequency. $T_{settling}$, including slew period, should be less than $T_s/2$=250ns.

This requirement for OTA is relatively easy, so it confirms that the second-order modulator is a satisfactory choice in this implementation.

Then we choose G1=G2=0.2 in Fig. 24, and more detailed circuit blocks are illustrated in Fig. 25.



(b) The circuit of the second order modulator



(b) the non-overlap clock used in circuit of (a)

Fig. 25: Circuit implementation of the second order modulator.

Where Cs1=60fF, Cf1=300fF, Cs21=90fF, Cs22=60fF, Cf2=300fF, which is determined based on G1 and G2.

OTA design

Figure 26 shows the OTA used in this implementation. The OTA is designed to be fully differential because this topology has several benefits. First of all, it doubles the effective output swing, whereas the output noise power only gets doubled, resulting in an SNR that is also doubled. Second, a fully differential OTA has better power supply rejection and better commode-mode rejection. Its drawback is the need for common mode feedback circuits.

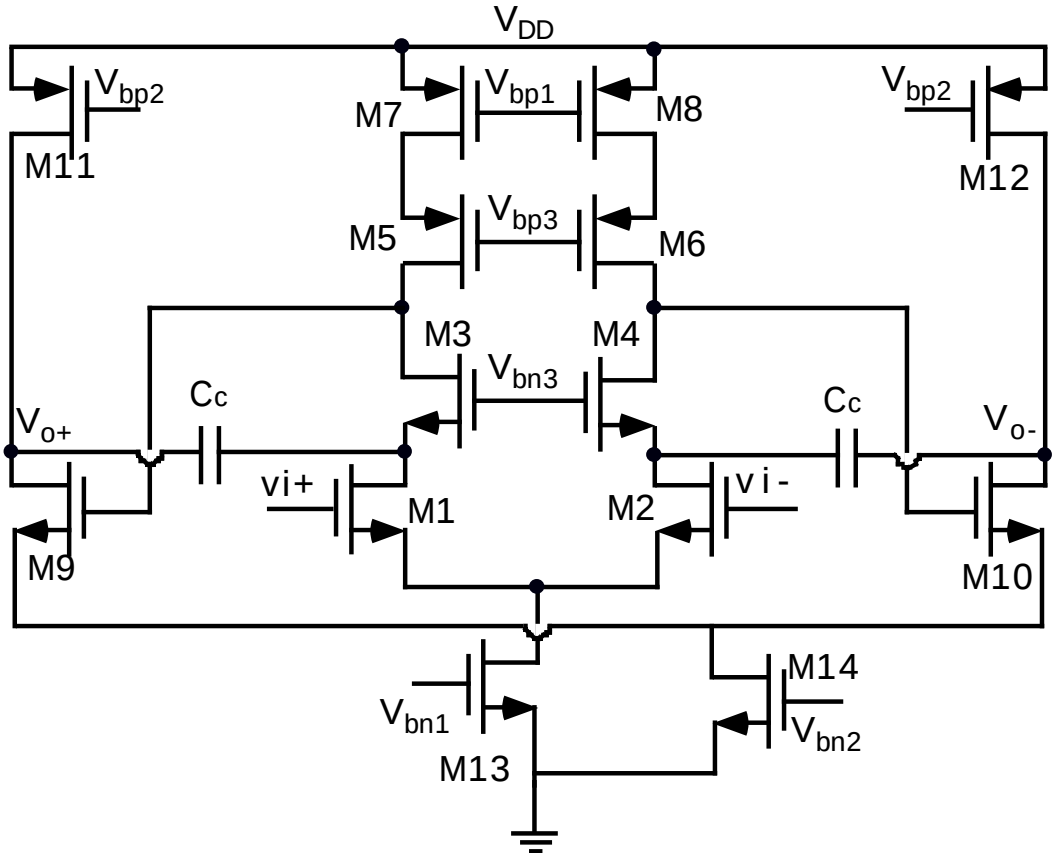


Fig. 26: Circuit diagram of the OTA.

The cascode compensation technique is used because this compensation technique was shown to provide a wider bandwidth than its Miller counterpart.

Biasing Circuit for the OTA:

The biasing for the OTA is shown in Fig. 27. Transistors Mb1-Mb4 construct current mirrors to set the bias voltage required by the OTA. The transistors forming the current mirror are chosen to have identical channel lengths to improve matching.

Common Mode Feedback:

Common-mode feedback is necessary to define the DC voltages at the high impedance output nodes. A dynamic common-mode feedback is selected because it saves power. Figure 28 illustrates the CMFB circuit used in the first gain stage.

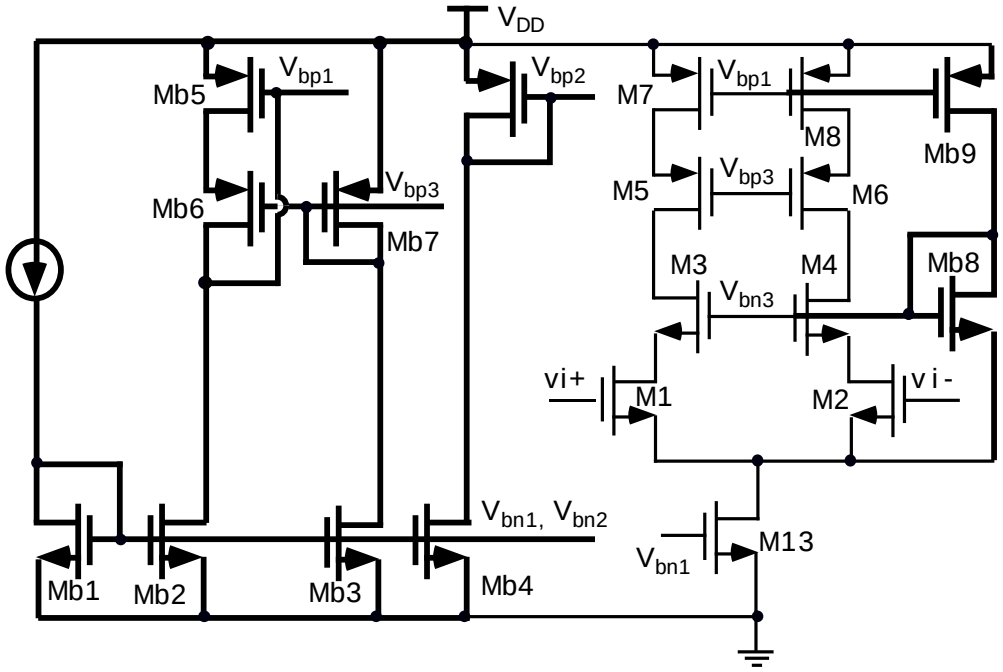


Fig. 27: Biasing circuit for the OTA.

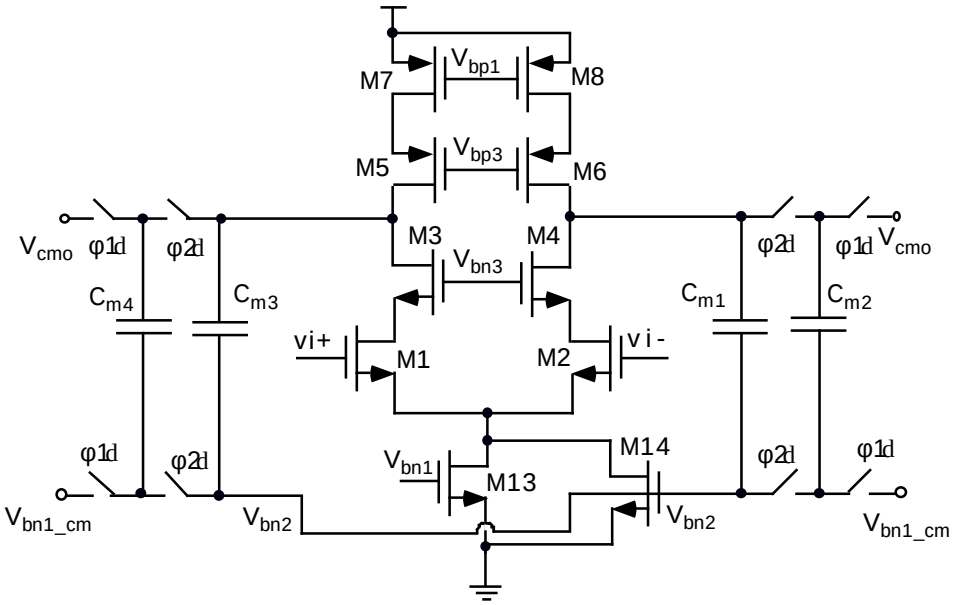


Fig. 28: Diagram for the common-mode feedback block for the fully-diffferential OTA.

This is a dynamic feedback, whose significant advantage is that it does not dissipate additional power. A similar network is used for the second gain stage, where Vcmo is the desired output common mode voltage, and Vbn1_cm has the same voltage as Vbn1 but is isolated from it, which can improve the stability of the circuit.

Simulation results:

Some of the simulation results of OTA are summarized in Table 3. The simulated frequency response and integration performance are shown in Figures 29 and 30.

Table 3: Summary of OTA performance.

Paremeters	*Specification*
Power consumption	1.688mW
Adc (with V_{od}=max)	120dB
Unit gain frequency	28.2MHz
Phase margin	84°
Settling time	120ns
Dynamic range	69dB

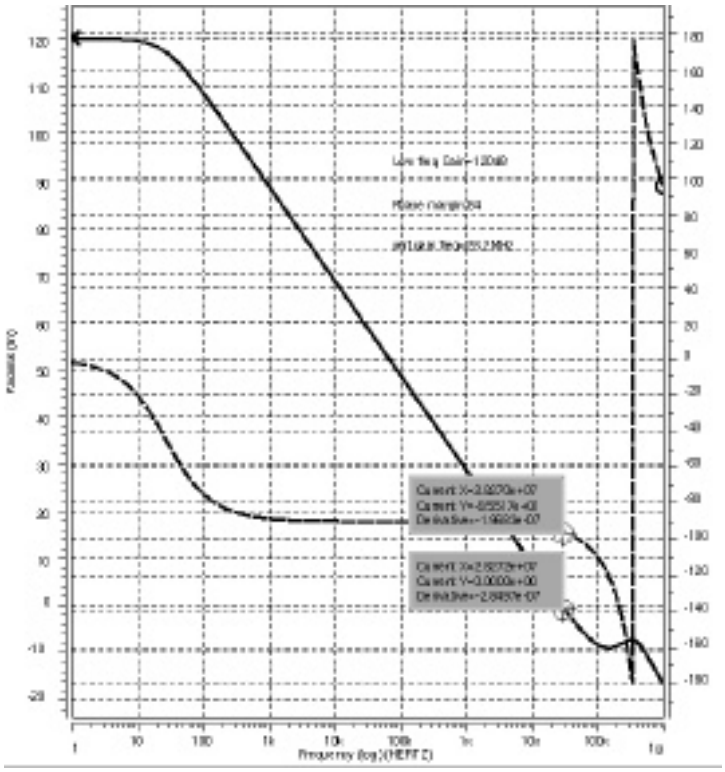


Fig. 29: AC performance of the OTA.

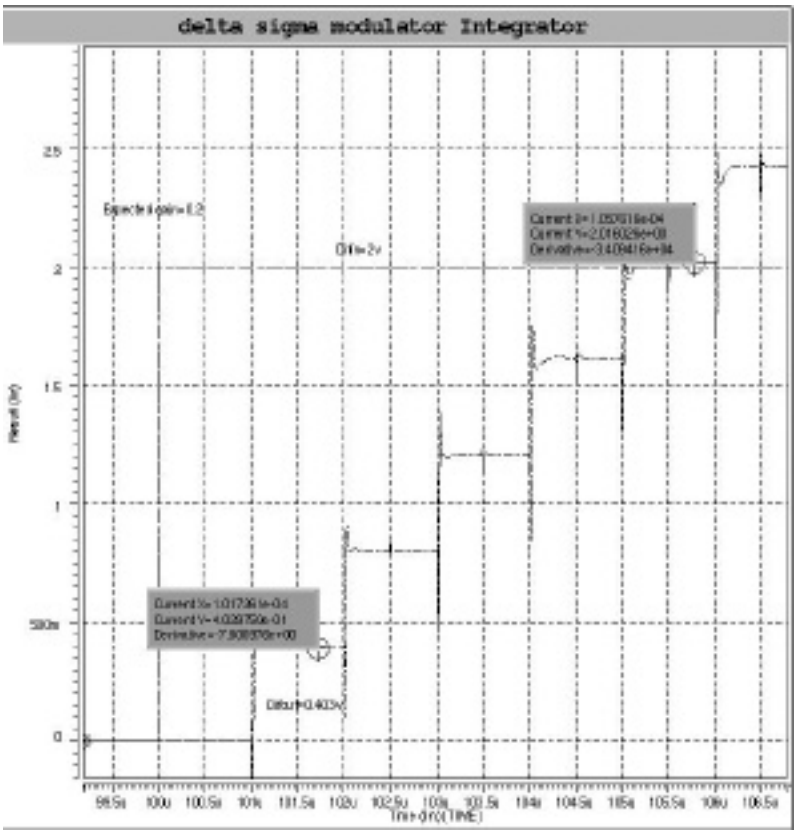


Fig. 30: Waveform of the differential output of the SC integrator.

5.2. Two-Phase Clock Generator

A two-phase non-overlapping clock (with delays) is required in the integrators in order to minimize signal dependent charge-injection errors. The waveform of the two-phase clock is shown in Fig. 25(b), where Φ_{1d}, Φ_{2d} are the delayed version of Φ_1 and Φ_2. The rising edges of the delayed and non-delayed clocks should be lined up with the rising so that the settling time for the OTA, given in equation, can be maximized.

$$t_{settling} = \frac{T_s}{2} - t_r - t_f - t_{nol}$$

where Ts is the sampling period
tnol is the non-overlap time
tr is the rise time
tf is the fall time

Figure 31 shows the circuit to realize two-phase clock.

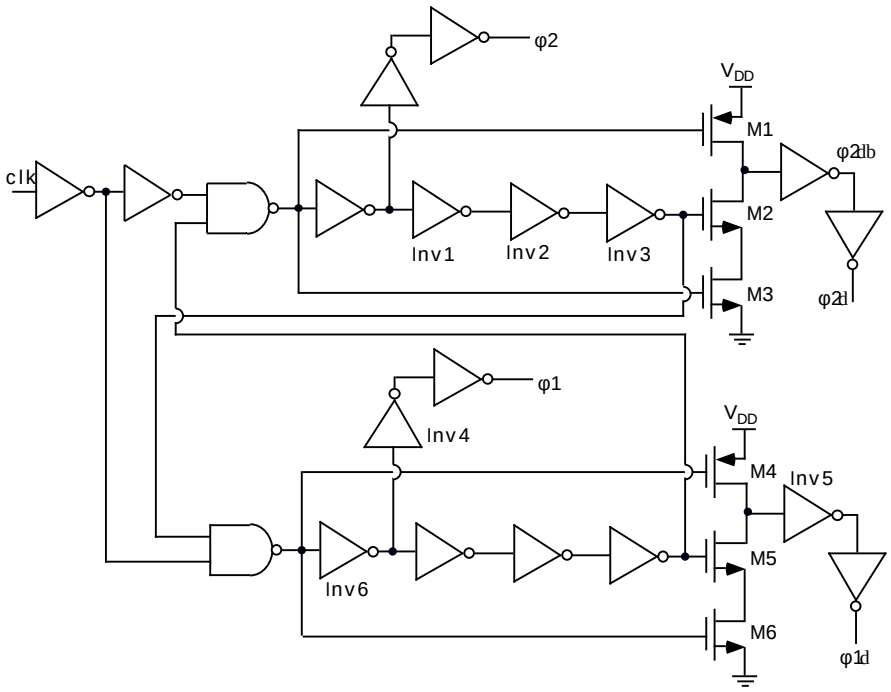


Fig. 31: The circuit schematic of the two-phase clock generator.

From the schematic, we can see that the delay time td1 between Φ_1 and Φ_{1d}, illustrated in Figure 1.2 (b), is determined by inverter Inv1, Inv2 and Inv3. Therefore, we can flexibly adjust t_d by sizing these inverters. Inverters Inv4 and Inv5 have the same sizing and M4's size is the same as pull-up PMOS in Inverter Inv6, so the rising edges of Φ_1 and Φ_{1d} are lined up. The same design strategy is used for Φ_2 and Φ_{2d}. Inverters Inv4 and Inv5 have the same sizing and M4's size is the same as pull-up PMOS in Inverter Inv6, so the rising edges of Φ_1 and Φ_{1d} are lined up. The same design strategy is used for Φ_2 and Φ_{2d}. Simulation results for this circuit are shown in Fig. 32.

The following parameters can be measured from the simulation:

$$tr \approx tf \approx 1.6ns,\ tnol=1.8ns,\ td=7ns$$

The power consumption of the clock generator is 226μW.

5.3. One-bit A/D Converter

The modulator also requires a one-bit A/D. This is a simple dynamic comparator, as shown in Fig. 33. When the Latch/Reset signal is low, the differential outputs are set to VDD. As it turns high, transistors M7 and M9 are cut off, whereas M5 and M6 begin

to conduct. The cross-coupled inverters, formed by M3, M4, M8 and M10, turn to the output, which is determined by the differential input signal at the rise edge of Latch/Reset.

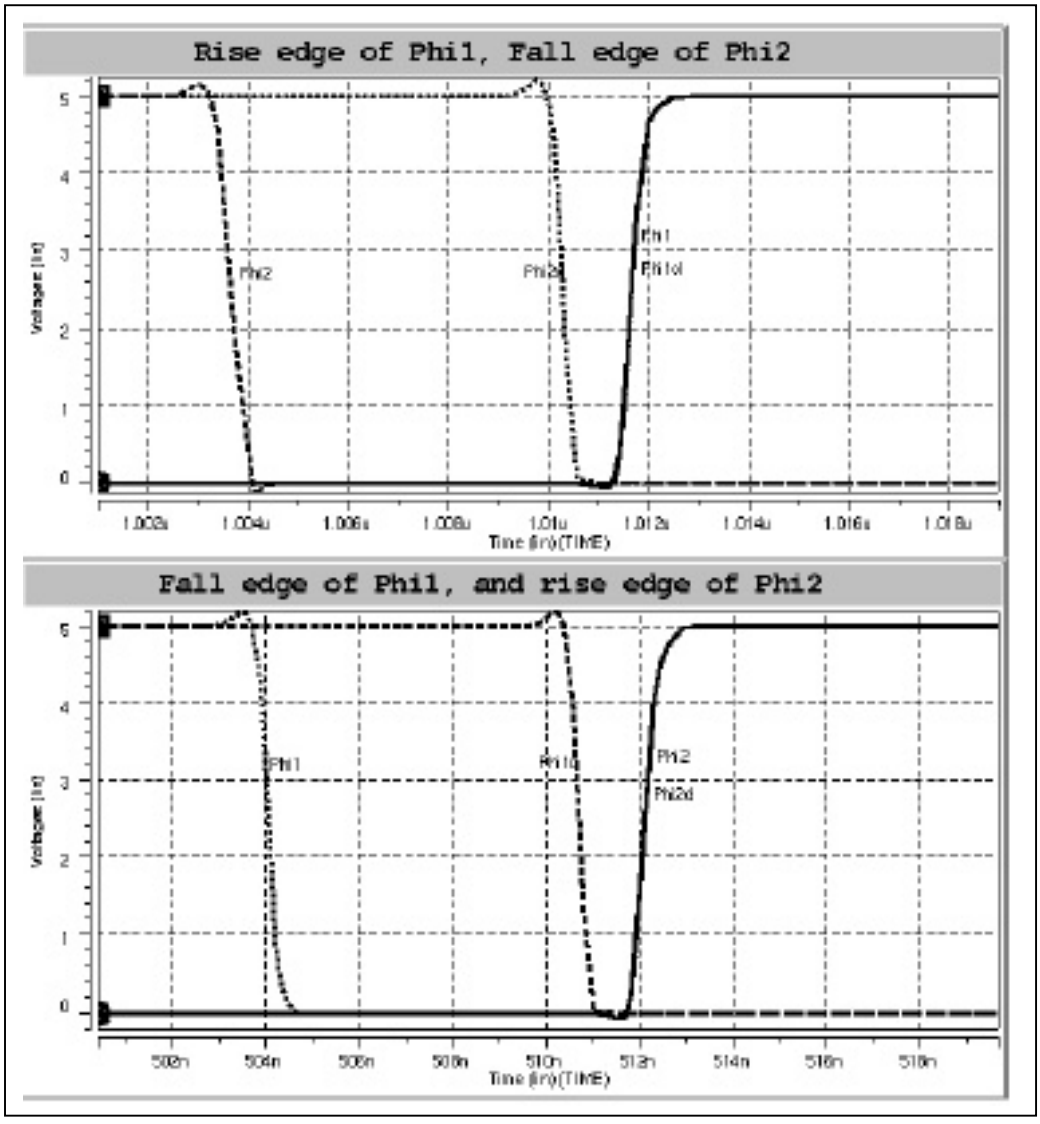


Fig. 32: Simulation of the two-phase clock generator.

The simulation results for this circuit are shown in Fig. 34. The simulation results are as expected. At the rising edge of the Latch signal, the comparator evaluates the input differential voltage. The output remains stable even when the input changes. When the Latch signal goes low, the outputs are reset to VDD and prepared for the next evaluation.

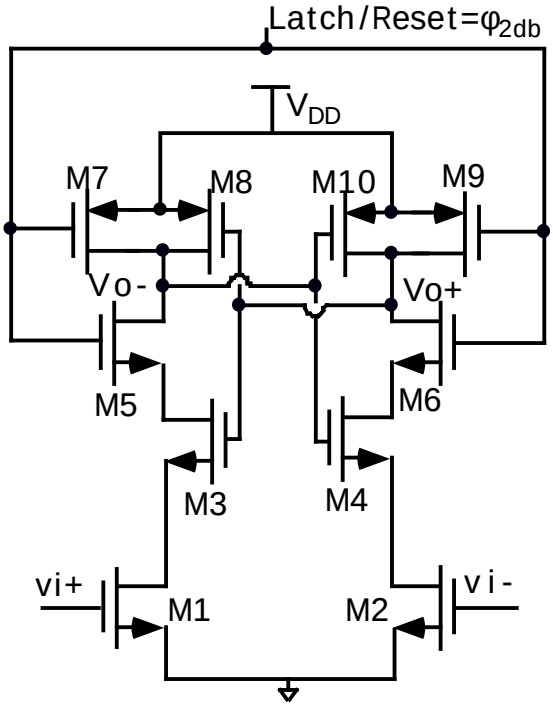


Fig. 33: Circuit diagram for the dynamic 1-bit ADC.

The power consumption of this block is 20μW for the latch signal with the clock frequency of 2.5MHz.

5.4. Design of the Decimator

The decimator transforms the input modulated signal from short words to long words, and it decreases the high sampling rate to Nyquist rate. The output of the modulator represents the input signal together with its out-of-band components, modulation noise, circuit noise and interference. A low pass filter is needed to attenuate all of the out-of-band noise of this signal, so when the modulated signal is re-sampled at the Nyquist rate, the aliasing noise penalty is insignificant.

An important criterion in the design of an A/D converter is the efficiency with which the decimator operation can be implemented. This efficiency is directly related to the type, the order and the architecture of the digital filter used in the implementation. Typically the order of a FIR low-pass filter is directly related to a function of the required ripples δ_p and δ_s in the pass-band and stop-band, respectively. It has the approximate form:

$$N \approx \frac{D_\infty(\delta_p, \delta_s)}{\Delta F / F}$$

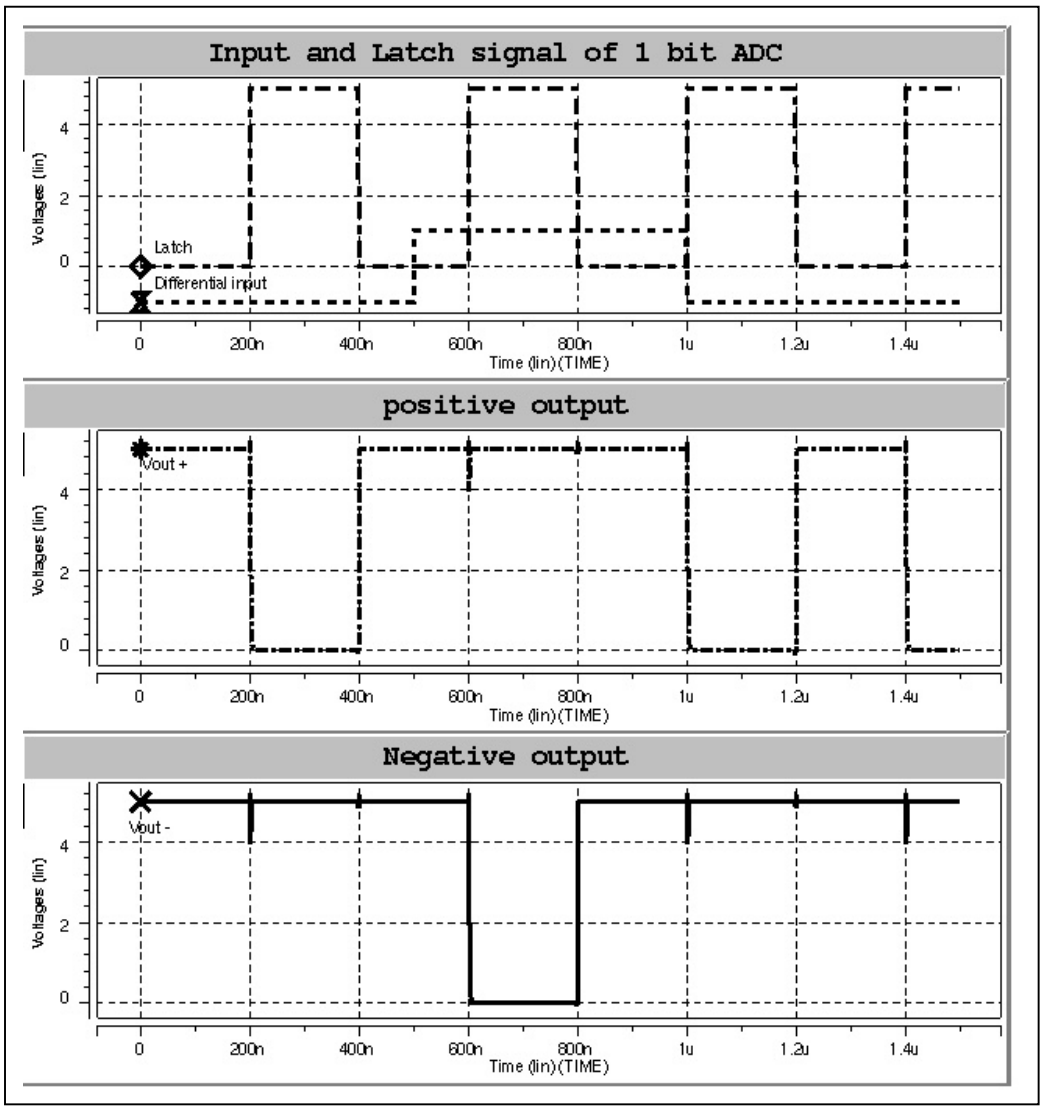


Fig. 34: Simulation results for the dynamic comparator.

where

$$D_\infty(\delta_p, \delta_s) = \log_{10}\delta_s\left[a_1(\log_{10}\delta_p)^2 + a_2\log_{10}\delta_p + a_3\right] + \left[a_4(\log_{10}\delta_p)^2 + a_5\log_{10}\delta_p + a_6\right]$$

and a1, a2, …, a6 are constants.

James C. Candy has shown that multistage decimators usually outperform single stage decimator in efficiency. Practical consideration leads to the conclusion that a two-stage design is the best choice. In addition, the choice of 2:1 for the last stage is both the

theoretical best option as well as the most practical one. A typical two-stage decimator diagram is shown in Fig. 35.

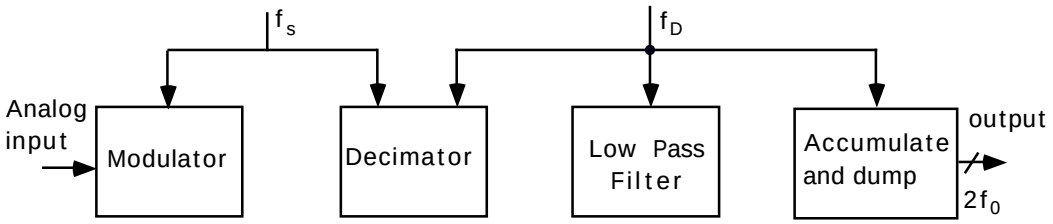


Fig. 35: The block diagram of two stage decimator

The first stage decimator, sinc^k decimator:

The First stage decimator is usually implemented by using sinc^k decimator. The frequency response of sinc^k decimator is given by

$$H(f) = [\frac{\sin c(\pi \quad fNT)}{\sin c(\pi \quad fT)}]^K$$

The decimation ratio N is

$$N = \frac{f_s}{f_D}$$

where f_s is the over-sampling frequency of the modulator, and f_D is the output frequency of the sinc^k decimator. It has been shown that a sinc^{k+1} filter is close to being optimum for decimating the signal from $\Delta\Sigma$ modulation of order of k. The penalty for using this class of decimation is typically less than a 0.5dB increase in noise. The order of our modulator is 2, so sinc^3 decimator is used in our implementation.

The Z transfer function of the filter is:

$$H(z) = [\frac{1 - z^{-N}}{1 - z^{-1}}]^3$$

The block diagram to implement the above transfer function is shown in Fig. 36.

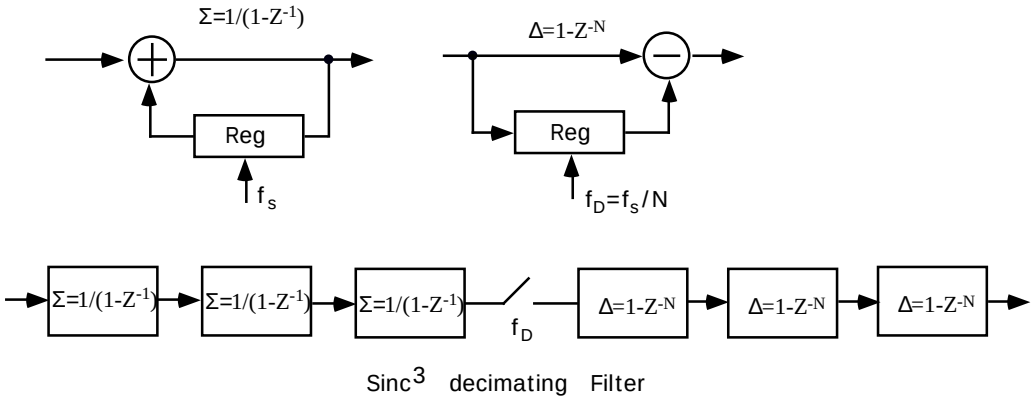


Fig. 36: The implementation of $Sinc^3$ decimation filter.

The second stage:

The low pass filter in the second stage is used for anti-aliasing; an FIR filter is chosen for our implementation. We design the cutoff frequency to be Nyquist rate of 20kHz. The accumulate and dump circuit in the second stage is shown in Fig. 37.

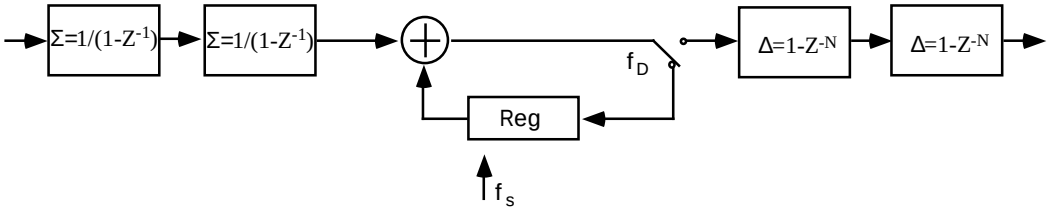


Fig. 37: The accumulate and dump circuit.

5.5. Layout of the Sigma-Delta Modulator

The above circuits for the modulator have been layed out. The layout plays an important role in ensuring that the expected performance is obtained. Some layout considerations are discussed below:

1) Matched transistors, capacitors and resistors: It is desirable that some transistors in the circuit design are exactly matched to optimize the circuit performance. This has been achieved by adopting a centroid layout where central symmetry is maintained throughout the circuit. This principle is also applied to layout matched capacitors and resistors.
2) Minimize the parasitic capacitance in high speed circuitry: We wish to minimize parasitic capacitance in the critical op-amps to increase bandwidth or reduce

settling time. Some layout skills, such as finger layout and waffle-iron layout techniques, are used in the transistor layout to achieve this goal.

3) Protect the input from the electrical surge or static charge: The circuitry is placed in a 40-pin pad frame. The input pins are connect to V_{DD} and GND through two diodes, where V_{DD} is connected to a n-terminal and GND is connected to a p-terminal respectively. Therefore, the input voltage variation can range from (GND-V_t)to (V_{DD}+V_t).

The layout is shown in Fig. 38.

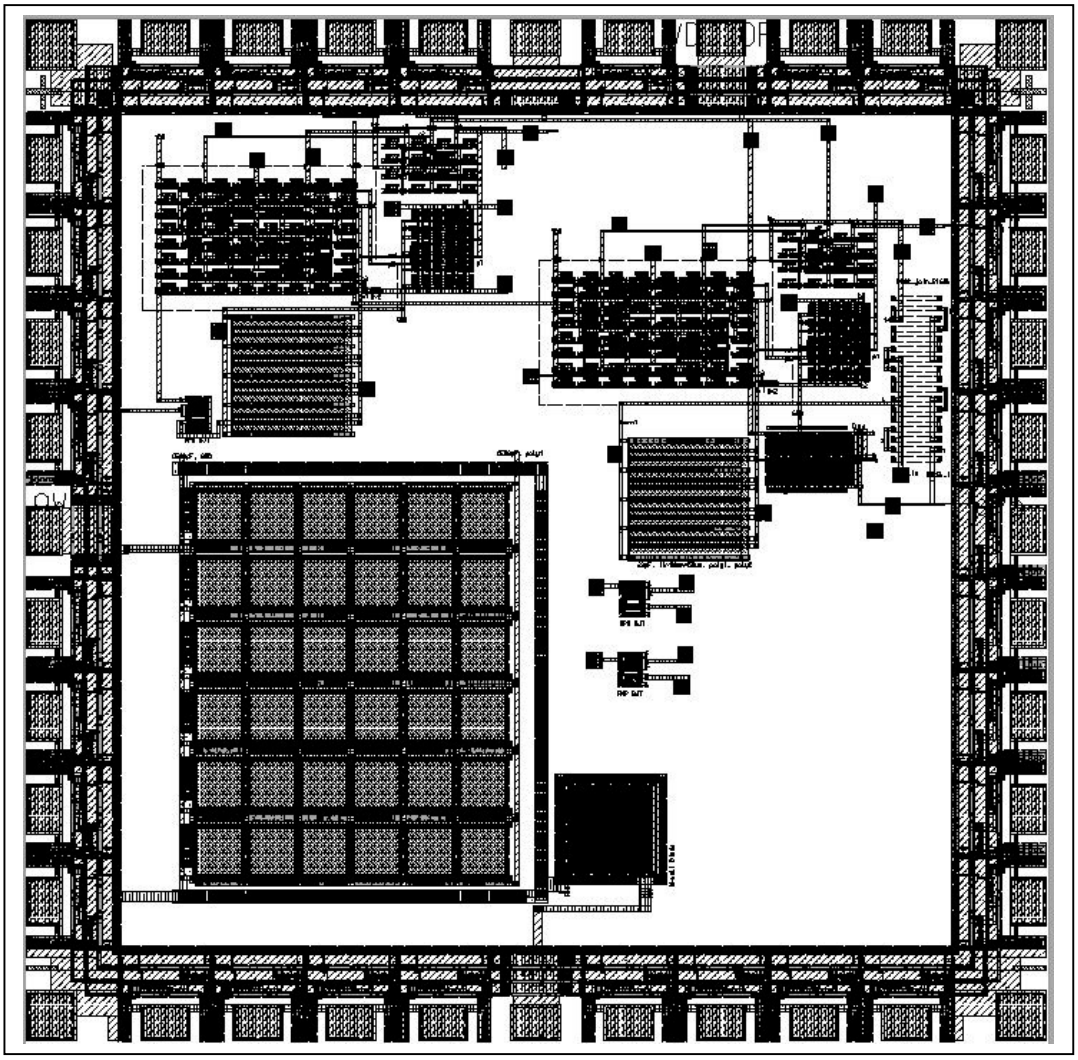


Fig. 38: The layout of the modulator and the revised regulator.

5.6. System Command and Data Transmission Protocol

As mentioned in previous reports, command and sensor data for the telemetry system will be communicated between the probes and the external world through a

wireless link. This past quarter work started to define and design the data transmission protocol, as summarized below.

1) External-to-internal transmission:

The external transmitter delivers power and commands to the implanted chip via an RF telemetry link. The method to modulate the transmitter carrier might be FSK (frequency shift keying) or ASK (amplitude shift keying). ASK is chosen in our application, because FSK circuitry is more difficult to implement than ASK circuitry as well as requiring tuning of the circuitry and more power consumption. The internal envelope detector, composed of a bandpass filter and a Schmitt trigger, is used to demodulate the ASK signals. Pulse width encoding is usually used in source coding of the external-to-internal transmission. Three different nominal envelopes, representing "0", "1", and synchronization bit, are shown in Fig. 39.

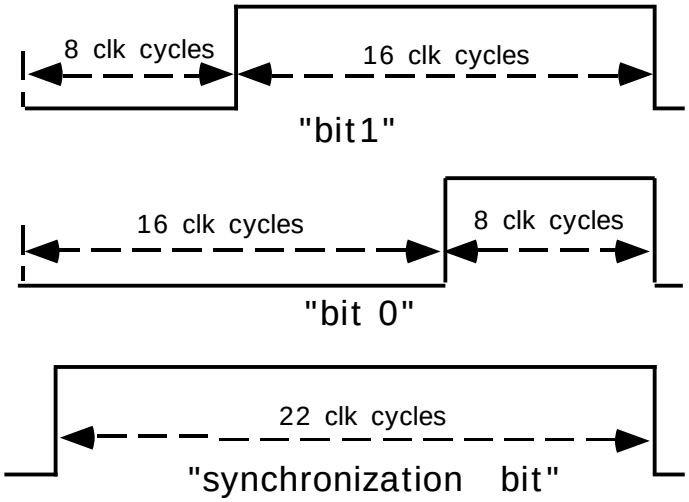


Fig. 39: Pulse width modulation for "0", "1", and synchronization bit.

It is obvious that a simple counter after the envelope detector can act as the decoder, which makes the logic circuitry design much easier. The transmission protocol is illustrated in Fig. 40.

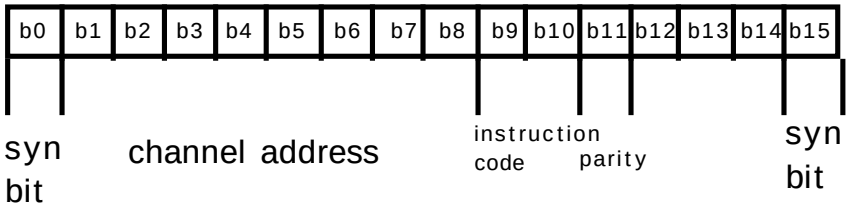


Fig. 40: External-to-Internal transmission protocol.

where b0 and b15 are synchronization bits used for communication synchronization.

b1-b8 are channel address codes, which can select from up to 256 channels.

b9-b10 are instruction code with definition as following,

00: standby
01: measure one channel
10: measure all channels in turn
11: measurement data acknowledgement

b11 is parity check bit
b12-b14 are reserved for future usage.

2) Internal-to-External transmission

The main purpose of internal-to-external transmitter is to deliver data to the external receiver. There are two methods to implement this function, one is to use an internal transmitter (active way), while the other (passive way) is to modulate the load of the implanted unit so that the external transmitter can sense this load change. The former method can achieve more reliable data communication but consumes more power, on the contrary, the latter method saves power but affects power delivery efficiency and the reliability of data transmission.

Based on past experience and measured data in our group, an active transmitter will be utilized to implement the internal-to-external transmission in this application. However, the passive telemetry technique implementation may be included as a reference.

For the same reasons as in (1), ASK is also chosen for internal-to-external data transmission. However, a specific ASK, OOK (On-off keying), can be used in this situation, because the off period can exist without the need to deliver power and it saves power. Pulse width encoding is also used in source coding of the internal-to-external transmission. The transmission protocol is illustrated in Fig. 41.

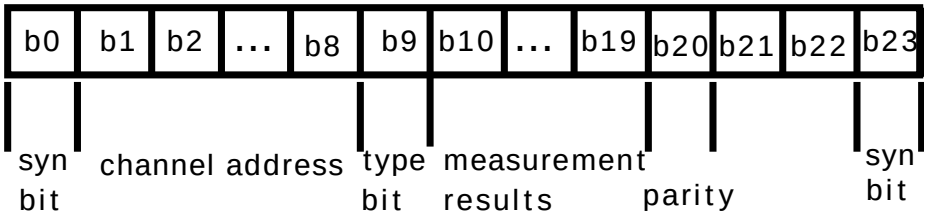


Fig. 41: The I-to-E transmission protocol.

where, b0 and b23 are synchronization bits used for communication synchronization.

b1-b8 are channel address codes, used as a handshake in reply.

b9: type bit
0: this signal is a command acknowledgement.
1: this signal is a measurement result.

b10-b19: measurement result.

b20: parity check bit.

b21-b22: reserved for future usage.

5.7. System Logic design

As shown in Fig. 42, the system logic is controlled by the central controller, which mainly accomplishes the following three functions.

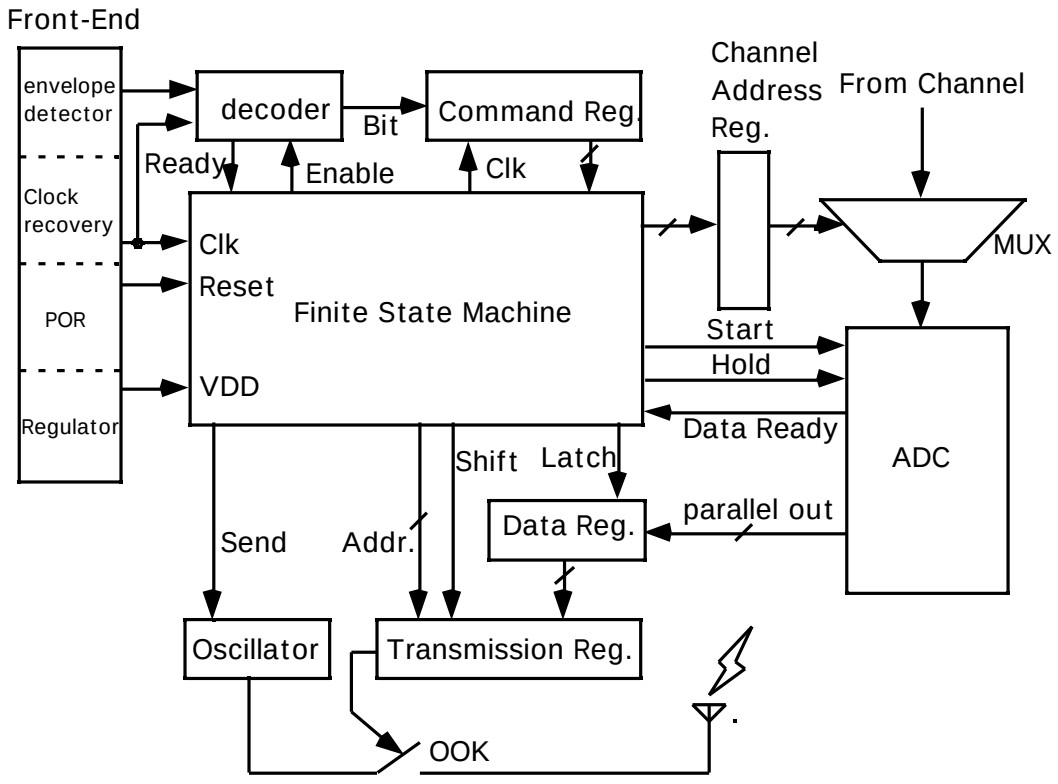


Fig. 42: Block diagram for overall system logic.

1. *Command decoding*
   The command decoder takes the signal from the envelope demodulator and obtains the command code by timing the pulse duration. Then the command bits are shifted to the 16-bit command register. After one complete command word is decoded correctly and saved, the decoder sends the ready signal to finite state machine.

2. *A/D control*
   First, the finite state machine latches the channel address into the address register, which controls the multiplexer to pass the selected channel signals to the A/D inputs. After the "start" signal is sent to initiate A to D conversion, the finite state machine waits until the "data ready" signal from ADC becomes effective, replying with "hold" to halt the ADC and fetching the digital data to the transmission register.

3. *Internal-to-external transmission control*
   For reliable communication, the central controller on the implanted chip is required to send acknowledge signals to the external receiver after correctly encoding every command word. It latches the channel address and necessary synchronization bits into the transmission register, then starts the oscillation of the internal transmitter. The bits in the transmission register are shifted to modulate the carrier by OOK (on-off-keying) mechanism.

   On the other hand, the digitized signal is also transmitted to the external receiver by the same mechanism described in the preceding paragraph. The only difference is that the data of interest from ADC are latched to the transmission register together with the channel address. Note that the external transmitter replies with acknowledgement to guarantee reliable recording as well.

The central controller incorporates the function s of circuit blocks on-chip to make the recording telemetry system a functional and flexible unit. Verilog is being used to design, simulate and verify the system logic, and the layout will be generated by the available CAD tools.

Plans For the Coming Quarter:

The design of the system logic, the transmitter, and receiver blocks are not complete and are still in progress. As mentioned above, a test chip containing the latest design of the front-end circuitry, and the sigma-delta modulator has been sent out to MOSIS for fabrication. The chip is expected back this quarter and will be tested. Based on these test results, a more complete chip containing the logic and internal active transmitter will be completed and sent out for fabrication. The goal is to have the first complete system ready for initial testing by the end of summer 2001.

6. Conclusions

During the past quarter, probes having 16 sites have been chronically implanted in guinea pig auditory cortex in order to study site lifetime as a function of site size and placement. The probes contained a 1000μm^2 site at the tip and a normal 177μm^2 site located in the center of the shank 25μm behind it. After 9 weeks, no significant differences in recording quality have been observed over these sites, although the impedances slowly fluctuate. We have also implanted 32-channel "brain-in-the-box" probes; recorded activity on the first of these implants stopped after two weeks, and we suspect a cable break. Another array has been implanted and is being monitored carefully.

We have also completed a fabrication run containing a large number of active probes. Fifteen different probe designs were fabricated, including 64- and 96-site structures. The front-end-selected Phoenix3 probe has been used to record spontaneous and driven activity from 64 sites on 8 shanks. Some of these active probes contain special test modes that allow a test signal to be generated on-chip and applied to the sites. This feature has been verified for the first time and found to work satisfactorily so that active probes can be tested in-vivo at any time to examine site impedances. The 96-site buffered probes have been used to record from single units in guinea pig auditory cortex and in cerebellum. They are also currently being used for experiments in rat hippocampus at Rutgers University.

The probes fabricated employ a variety of readout circuits including capacitively-coupled probes designed to avoid dc offsets. These probes have measured gains of 36dB, bandwidths from 68Hz to 25kHz, and a power dissipation of 77μW. An eight-channel multiplexed probe has also been fabricated. Optically-induced offsets remain a problem on some of these probes, and the circuitry is being redesigned to eliminate the problem. In an effort to increase the number of sites that can be monitored, an in-vivo spike-detection circuit is being designed. It will be fabricated using the HP 0.5μm process available through MOSIS and mounted on the platform holding the probes. The spike detection threshold an be input by the user so that the location and time of occurrence of the spikes will be transmitted, eliminating the bandwidth consumed in transmitting baseline noise in full-analog systems. Eventually, the digitized spike amplitude could also be transmitted. The system is being designed to handle 1024 sites in a 3D array with 128 active channels. At a 2.56MHz clock, the time of any spike can be resolved to less than 0.4μsec.

Work has also gone forward with the design of a wireless interface for the probes. A 10b sigma-delta analog-to-digital converter (ADC) has been designed to achieve a Nyquist sampling rate of 20kHz. The dynamic range is 62dB. Simulations of the operational transconductance amplifier show a power dissipation of 1.7mW, a dc gain of 120dB, a unity-gain frequency of 28MHz and a dynamic range of 69dB. A two-phase non-overlapping clock has also been designed for this system with a simulated power dissipation of 226μW. A decimator has also been designed. The ADC has been laid out and signal and data transmission protocols for the wireless system have been defined.

The system will use an amplitude-shift keyed format. The goal is a fully wireless recording system by the end of summer.